SI. No	Title	Page no
5.2	Scaling Techniques	1-5
5.2.1	Comparative Scaling Techniques	6-12
5.2.2	Non Comparative Scaling Techniques	13-25



#### 5.2 Scaling – Meaning

Scaling is a process or set of procedures, which is used to assess the attitude of an individual. Scaling is defined as the assignment of objects to numbers according to a rule. The objects in the definition are text statements, which can be the statements of attitude or principle. Attitude of an individual is not measured directly by scaling. It is first migrated to statements and then the numbers are assigned to them.

Here a respondent is asked to show his preferences from among five brands of coffee -A, B, C, D and E with respect to flavours. He is required to indicate his preference in pairs. A number of pairs are calculated as follows. The brands to be rated are presented two at a time, so each brand in the category is compared once to every other brand. In each pair, the respondents were asked to divide 100 points on the basis of how much they liked one compared to the other. The score is totally for each brand.

The researcher begins with the development of **construct of interest**. The **construct** refers to the type of concept existing at the high level of abstraction than do everyday concepts, such as brand loyalty, attitude, product involvement, satisfaction, etc.

Next, the researcher develops the **theoretical definition of the construct** that clearly states or defines the concept or central idea of interest. In order to do so, one needs to have an underlying theory of the construct that can be measured. A defined theory is required not only for the construction of a scale but also for the interpretation of the resulting data. **For example,** brand loyalty can be advocated by a consistent



#### Module 5

### Part 2 Scaling Techniques Dr Josheena Jose, Assistant Professor, Christ College (Autonomous) Irinjalakuda

repurchase of a brand as a result of a respondent's favorable attitude towards the brand.

The construct must be **operationalized** in such a way that it complies with the theoretical definition. The **operational definition** implies that which observable characteristics or attributes will be measured and how the numerical score is assigned to each. **For example,** in the context of tea, a consumer will be characterized as brand loyal if he shows a favorable attitude and must have purchased the same brand of tea for at least five out of last six purchase occasions.

In the next step, the **initial pool of scale items is generated** on the basis of theory, analysis of the secondary data and qualitative research. Out of this, a reduced set of prospective scale items is generated based on a qualitative judgment of a researcher and other expert individuals. The reduced scale items are further reduced in a quantitative manner since these are still too large to constitute a scale.

Once the reduced prospective set of scale items is obtained, the **researcher collects the data** on it from a large pretest sample of respondents. The data obtained is analyzed using the statistical techniques such as correlation, cluster analysis, confirmatory factor analysis, statistical tests, exploratory factor analysis and discriminant analysis.

As a result of the statistical analysis, further more items are eliminated, thereby resulting in a **purified scale**. The purified scale is then evaluated for reliability, validity and generalizability by **collecting more data from different samples**. As a result of such assessment, **a final set of scale items is selected**.



#### Module 5 Part 2 Scaling Techniques Dr Josheena Jose, Assistant Professor, Christ College (Autonomous) Irinjalakuda A paradigm for its development is shown below:



#### **Other Scaling Techniques**

Scaling of objects can be used for a comparative study between more than one objects (products, services, brands, events, etc.). or can be individually carried out to understand the consumer's behaviour and response towards a particular object.

Following are the two categories under which other scaling techniques are placed based on their comparability:



• **Comparative scales** involve the direct comparison of stimulus objects. Comparative scale data must be interpreted in relative terms and have only ordinal or rank order properties.

In **non comparative scales**, each object is scaled independently of the others in the stimulus set. The resulting data are generally assumed to be interval or ratio scaled.





#### **Comparative Scales**

For comparing two or more variables, a comparative scale is used by the respondents. Following are the different types of comparative scaling techniques:

# **4** Paired Comparison Scaling

**Definition:** The **Paired Comparison Scaling** is a comparative scaling technique wherein the respondent is shown two objects at the same time and is asked to select one according to the defined criterion. **The resulting data are ordinal in nature.** 

The paired Comparison scaling is often used when the stimulus objects are physical products. The comparison data so obtained can be analyzed in either of the ways. First, the researcher can compute the **percentage of respondents** who prefer one object over another by adding the matrices for each respondent, dividing the sum by the number of respondents and then multiplying it by 100. Through this method, all the stimulus objects can be evaluated simultaneously.

Second, under the **assumption of transitivity** (which implies that if brand X is preferred to Brand Y, and brand Y to brand Z, then brand X is preferred to brand Z) the paired comparison data can be converted into a rank order. To determine the rank order, the researcher identifies the number of times the object is preferred by adding up all the matrices.

The paired comparison method is effective when the number of objects is limited because it requires the direct comparison. And with a large number of stimulus objects the comparison becomes cumbersome. Also, if there is a violation of the



#### Module 5

## Part 2 Scaling Techniques Dr Josheena Jose, Assistant Professor, Christ College (Autonomous) Irinjalakuda

assumption of transitivity the order in which the objects are placed may bias the results.

Example: Here a respondent is asked to show his preferences from among five brands of coffee – A, B, C, D and E with respect to flavours. He is required to indicate his preference in pairs. A number of pairs are calculated as follows. The brands to be rated are presented two at a time, so each brand in the category is compared once to every other brand. In each pair, the respondents were asked to divide 100 points on the basis of how much they liked one compared to the other. The score is totally for each brand. No. of pairs = N(N 1) /2

In this case, it is 5(5-1)/2

#### A&B B&D A&C B&E A&D C&D A&E C&E B&C D&E

A paired comparison symbolizes two variables from which the respondent needs to select one. This technique is mainly used at the time of product testing, to facilitate the consumers with a comparative analysis of the two major products in the market.

To compare more than two objects say comparing P, Q and R, one can first compare P with Q and then the superior one (i.e., one with a higher percentage) with R.

*For example*, A market survey was conducted to find out consumer's preference for the network service provider brands, A and B. The outcome of the survey was as follows:



Brand 'A' = 57%

Brand 'B' = 43%

Thus, it is visible that the consumers prefer brand 'A', over brand 'B'.

kank Order Scaling

**Definition:** The **Rank Order Scaling** is a yet another comparative scaling technique wherein the respondents are presented with numerous objects simultaneously and are required to order or rank these according to some specified criterion.

The Rank order scaling is often used to measure the preference for the brand and attributes. The ranking data is typically obtained from respondents in the **conjoint analysis** (*a statistical technique used to determine how the brand and the combination of its attributes such as features, functions, and benefits, influences the decision making of a person*), as it forces the respondents to discriminate among the stimulus objects. The Rank order scaling results in the ordinal data.

With respect to the paired comparison scaling, the Rank order scaling resembles more closely to the shopping environment, and also it takes less time and eliminates all the intransitive responses (not object-directed). Such as, if there are 'n' stimulus objects, then only 'n-1' scaling decisions are to be made in case of Rank order scaling, while in the case of paired comparison scaling '[n (n-1) /2]' scaling decisions are required.



Moreover, the rank order scaling is an easy method to understand. But, however, the major limitation of this process is that it results only in ordinal data.

**Note:** Under the assumption of **Transitivity** (*implies that if brand X is preferred to brand Y, and brand Y, is preferred to brand Z, then brand X is preferred to brand Z*), the rank order data can be converted to equivalent paired comparison data and vice-versa.

In rank order scaling the respondent needs to rank or arrange the given objects according to his or her preference.

*For example*, A soap manufacturing company conducted a rank order scaling to find out the orderly preference of the consumers. It asked the respondents to rank the following brands in the sequence of their choice:

SOAP BRANDS	RANK
Brand V	4
Brand X	2
Brand Y	1
Brand Z	3



The above scaling shows that soap 'Y' is the most preferred brand, followed by soap 'X', then soap 'Z' and the least preferred one is the soap 'V'.

#### 븆 Constant Sum Scaling

**Constant Sum Scaling** 

**Definition:** The **Constant Sum Scaling** is a technique wherein the respondents are asked to allocate a constant sum of units, such as points, dollars, chips or chits among the stimulus objects according to some specified criterion.

In other words, a scaling technique that involves the assignment of a fixed number of units to each attribute of the object, reflecting the importance a respondent attaches to it, is called as constant sum scaling. **For example**, Suppose a respondent is asked to allocate 100 points to the attributes of a body wash on the basis of the importance he attaches to each attribute. In case he feels any attribute being unimportant can allocate zero points and in case some attribute is twice as important as any other attribute can assign it twice the points. The sum of all the points allocated to each attribute should be equal to 100.

Once the points are allocated, the attributes are scaled by counting the points as assigned by the respondents to each attribute and then dividing it by a number of respondents under analysis. Such type of information cannot be obtained from rank order data unless it is transformed into interval data. The constant sum scaling is considered as an ordinal scale because of its comparative nature and lack of generalization.



One of the advantages of the constant sum scaling technique is that it allows a proper discrimination among the stimulus objects without consuming too much time. But however, it suffers from two serious limitations. First, the respondent might allocate more or fewer units than those specified. Second, there might be a rounding error, in case too few units are allocated. On the other hand, if a large number of units are used then it might be burdensome on the respondents and causes confusion and fatigue.

*For example*, The respondents belonging to 3 different segments were asked to allocate 50 points to the following attributes of a cosmetic product 'P':

ATTRIBUTES	SEGMENT 1	SEGMENT 2	SEGMENT 3
Finish	11	8	9
Skin Friendly	11	12	12
Fragrance	7	11	8
Packaging	9	8	10
Price	12	11	11
Total	50	50	50

From the above constant sum scaling analysis, we can see that:



- Segment 1 considers product 'P' due to its competitive price as a major factor.
- But segment 2 and segment 3, prefers the product because it is skin-friendly.

# 4 Q-Sort Scaling

**Definition:** The **Q-Sort Scaling** is a Rank order scaling technique wherein the respondents are asked to sort the presented objects into piles based on similarity according to a specified criterion such as preference, attitude, perception, etc.

In other words, a scaling technique in which the respondents sort the number of statements or attitudes into piles, usually of 11, on the basis of some specified criterion. For example, suppose the respondents are given 100 motivational statements on individual cards and are asked to place these in 11 piles, ranging from the "most agreed with" to the "least agreed with". Generally, the most agreed statement is placed on the top while the least agreed statement at the bottom. The Q-Sort Scaling was developed to facilitate a quick discrimination among a relatively large number of stimulus objects. Thus, Q-Sort Scaling helps in assigning ranks to different objects within the same group, and the differences among the groups (piles) is clearly visible.

Q-sort scaling is a technique used for sorting the most appropriate objects out of a large number of given variables. It emphasizes on the ranking of the given objects in a descending order to form similar piles based on specific attributes.

• It is suitable in the case where the number of objects is not less than 60 and more than 140, the most appropriate of all ranging between 60 to 90.



*For example*, The marketing manager of a garment manufacturing company sorts the most efficient marketing executives based on their past performance, sales revenue generation, dedication and growth.

The Q-sort scaling was performed on 60 executives, and the marketing head creates three piles based on their efficiency as follows:



In the above diagram, the initials of the employees are used to denote their names.

#### **Non-Comparative Scales**

A non-comparative scale is used to analyse the performance of an individual product or object on different parameters. Following are some of its most common types:

#### **Continuous Rating Scales**

The **Continuous Rating Scale** is a Non comparative Scale technique wherein the respondents are asked to rate the stimulus objects by placing a point/mark appropriately on a line running from one extreme of the criterion to the other variable criterion.



It is a graphical rating scale where the respondents are free to place the object at a position of their choice. It is done by selecting and marking a point along the vertical or horizontal line which ranges between two extreme criteria.

*For example*, A mattress manufacturing company used a continuous rating scale to find out the level of customer satisfaction for its new comfy bedding. The response can be taken in the following different ways (stated as versions here):



The above diagram shows a non-comparative analysis of one particular product, i.e. comfy bedding. Thus, making it very clear that the customers are quite satisfied with the product and its features.

#### **Itemized Rating Scale**

The **Itemized Rating Scale** is an Ordinal Scale that has a brief description or numbers associated with each category, ordered in terms of scale positions. The respondents are asked to select the category that best describes the stimulus object being rated.



Itemized scale is another essential technique under the non-comparative scales. It emphasizes on choosing a particular category among the various given categories by the respondents. Each class is briefly defined by the researchers to facilitate such selection.

The itemized rating scale is widely used in marketing research and serve as a basic component of more complex scales, such as Multi-Item Scales. The three most commonly used itemized rating scales are as follows:

- Likert Scale: In the Likert scale, the researcher provides some statements and ask the respondents to mark their level of agreement or disagreement over these statements by selecting any one of the options from the five given alternatives. *For example*, A shoes manufacturing company adopted the Likert scale technique for its new sports shoe range named Z sports shoes. The purpose is to know the agreement or disagreement of the respondents. For this, the researcher asked the respondents to circle a number representing the most suitable answer according to them, in the following representation:
  - 1 Strongly Disagree
  - 2 Disagree
  - 3 Neither Agree Nor Disagree
  - 4 Agree
  - 5 Strongly Agree



STATEMENT	STRONGLY DISAGREE	DISAGREE	NEITHER AGREE NOR DISAGREE	AGREE	STRONGLY AGREE
Z sports shoes are very light weight	1	2	3	4	5
Z sports shoes are extremely comfortable	1	2	3	4	5
Z sports shoes look too trendy	1	2	3	4	5
I will definitely recommend Z sports shoes to friends, family and colleagues	1	2	3	4	5

The above illustration will help the company to understand what the customers think about its products. Also, whether there is any need for improvement or not.



• Semantic Differential Scale: A bi-polar seven-point non-comparative rating scale is where the respondent can mark on any of the seven points for each given attribute of the object as per personal choice. Thus, depicting the respondent's attitude or perception towards the object.

*For example*, A well-known brand for watches, carried out semantic differential scaling to understand the customer's attitude towards its product. The pictorial representation of this technique is as follows:

	Semantic Differential Scale													
	+3		+2		+1		0		-1		-2		-3	
Stylish .		:	_	:		:		:		:	X	:		Unfashionable
Simple .		:	_	:	x	:	_	:		:		:		Complex
Affordable.		:	_	:		:		:	Х	:	_	:		Expensive
High Quality	x	:		:		:		:		:		:		Low Quality
Nide Variety		:		:		:	X	:		:	į	:		Limited Variety

From the above diagram, we can analyze that the customer finds the product of superior quality; however, the brand needs to focus more on the styling of its watches. Sometimes the negative adjectives are placed on the right and sometimes on the left side of a scale. This is done to control the tendency of the respondents, especially those with either very positive or negative attitudes, to mark the right or left-hand sides of a scale without reading the labels.



The items on a semantic differential scale can be scored on either a numerical range of -3 to +3 or 1 to 7. The data obtained are analyzed through **profile analysis**. In profile analysis, the means and medians of the scale values are found out and then are compared by plotting or statistical analysis. Through this method, it is possible to compare the overall similarities and differences among the objects.

The versatility of the semantic differential scale increases its application in the marketing research. It is widely used in comparing the brand, company image, and product. It also helps in developing an advertising campaign and promotional strategies in new product development studies.

#### **Stapel Scale**:

**Definition: Stapel Scale** is a unipolar (one adjective) itemized rating scale designed to measure the respondent's attitude towards the object or event. The scale is comprised of **10 categories** ranging from -5 to +5 without any **neutral point** (zero).

The stapel scale got its name after its developer **Jan Stapel**. The scale is usually constructed vertically with a single adjective in the middle of the range of values (-5 to +5). The respondent is asked to select the appropriate numerical response category that best describes the extent to which the adjective related to the object is accurate or inaccurate. The higher the positive score selected by the respondent, the more accurate the adjective describes the object and vice versa.



# Module 5 Part 2 Scaling Techniques Dr Josheena Jose, Assistant Professor, Christ College (Autonomous) Irinjalakuda For example, the respondent is asked to rank the quality of food, and crew member

service of an airline on a scale ranging from -5 to +5:

+5	+5
+4 X	+4
+3	+3 X
+2	+2
+1	+1
High Food Quality	Poor Crew Service
-1	-1
-2	-2
-3	-3
-4	-4
-5	-5

From the above example, the airline is evaluated as having a high food quality but somewhat a poor cabin crew service.

The stapel scale is like a semantic differential scale with little modifications. It is often used in the situations when two bi-polar adjectives are difficult to find out. The data obtained are the interval and are analyzed in the same manner as the semantic differential data.

One of the advantages of the stapel scale is that it does not require the pre-testing of adjectives to ensure their true bipolarity. Also, it can be easily administered over the



telephone. But, however, some researcher finds the stapel scale to be quite complex and confusing.

#### **Multi-Item Scale**

**Definition:** A **multi-item scale** is designed to measure the respondent's attitude towards more than one attribute related to the stimulus object. It is a rating scale comprising of multiple items, where each item is a single question or a statement that needs to be evaluated.

#### **Other scaling Techniques**

#### **Thurstone Scale**

This is also known as an equal appearing interval scale. The following are the steps to construct a Thurstone Scale:

Step 1: To generate a large number of statements, relating to the attitude to be measured.

Step 2: These statements (75 to 100) are given to a group of judges, say 20 to 30, who were asked to classify them according to the degree of favourableness and unfavourableness.

Step 3: 11 piles are to be made by the judges. The piles vary from "most unfavourable" in pile 1 to neutral in pile 6 and most favourable statement in pile 11.



Step 4: Study the frequency distribution of ratings for each statement and eliminate those statements, which different judges have given widely scattered ratings.

Step 5: Select one or two statements from each of the 11 piles for the final scale. List the selected statements in random order to form the scale.

Step 6: The respondents whose attitudes are to be scaled were given the list of statements and asked to indicate their agreement or disagreement with each statement. Some may agree with one statement while some may agree with more than one statement.

Example

Suppose, we are interested in the attitude of certain socio-economic class of respondents towards savings and investments. The final list of statements would be as follows:

(a) One should live for the present and not the future. So, savings are absolutely not required.

(b) There are many attractions to spend the money saved.

(c) It is better to spend savings than risk them in investments.

(d) Investments are unsafe as the money is also blocked.

(e) You earn to spend and not to invest.

(f) It is not possible to save these days.

(g) A certain amount of income should be saved and invested.

(h) The future is uncertain and investments will protect us.



(i) Some amount of savings and investments are a must for every individual.

(j) One should try to save more so that most of it could be invested.

(k) All savings should be invested for the future.

Conclusion: A respondent agreeing to statements 8, 9 and 11 would be considered having a favourable attitude towards savings and investments. The person agreeing with statements 2, 3 and 4 is an individual with an unfavourable attitude. Also, if a respondent chooses statements 1, 3, 7 or 9, his attitude is not considered consistent.

#### Guttman scale

In the social sciences, the Guttman scale is often used to measure an increasing amount of "attitude" towards a single topic. In **education**, it's sometimes used to show a student's logical progression through coursework. Also known as Scalograms –cumulative scale. Both the respondents and items are ranked. The respondents total score is calculated by counting the No. of items answered favourably.

For example, the expected progression through math topics for 3 children. It's expected that a child does well in fractions before they are able to grasp algebra. A "0" means that the student hasn't mastered a topic, while a "1" means that a student has mastered it:



Problems										
Addition	Fractions	Bar graphs	Algebra							
1	0	0	0							
1	1	1	0							
1	1	0	1							
	Addition 1 1 1	Problems Addition Fractions 1 0 1 1 1 1 1 1	ProblemsAdditionFractionsBar graphs100111110							

In practice, it's rare to find data that fits perfectly to a Guttman scale. More often than not, you're actually testing more than one factor.

#### **Balanced or Unbalanced Rating Scale?**

Balanced rating scale: A fixed-alternative rating scale with an equal number of positive and negative categories; a neutral point or point of indifference is at the center of the scale.

The fixed-alternative format may be balanced or unbalanced. For example, the following question, which asks about parent-child decisions relating to television program watching, is a balanced rating scale:

Who decides which television programs your children watch?

- 4 Child decides all of the time.
- 4 Child decides most of the time.
- 4 Child and parent decide together.
- 4 Parent decides most of the time.
- 4 Parent decides all of the time.



This scale is balanced because a neutral point, or point of indifference, is at the center of the scale.

**unbalanced rating scale:** A fixed-alternative rating scale that has more response categories at one end than the other, resulting in an unequal number of positive and negative categories.

**Unbalanced rating scales** may be used when responses are expected to be distributed at one end of the scale. Unbalanced scales, such as the following one, may eliminate this type of "end piling":

- 4 Completely Dissatisfied
- ∔ Dissatisfied
- 🖊 Somewhat Dissatisfied
- ∔ Satisfied
- Completely Satisfied

Notice that there **are three "satisfied" responses and only two "dissatisfied"** responses above. The choice of a balanced or unbalanced scale generally depends on the nature of the concept or the researcher's knowledge about attitudes toward the stimulus to be measured.

**Odd or even number of categories** With an odd number of categories, the middle scale position is generally designated as neutral or impartial. The presence, position and labelling of a neutral category can have a significant influence on the response. The Likert scale is a balanced rating scale with an odd number of categories and a neutral point. The decision to use an odd or even number of categories depends on whether some of the participants may be neutral on the response being measured. If a neutral or indifferent response is possible from at least some of the participants, an



odd number of categories should be used. If, on the other hand, the researcher wants to force a response or believes that no neutral or indifferent response exists, a rating scale with an even number of categories should be used.

#### Forced versus non-forced choice

On forced rating scales the participants are forced to express an opinion because a 'no opinion' option is not provided. In such a case, participants without an opinion may mark the middle scale position. If a sufficient proportion of the participants do not have opinions on the topic, marking the middle position will distort measures of central tendency and variance. In situations where the participants are expected to have no opinion, as opposed to simply being reluctant to disclose it, the accuracy of data may be improved by a non-forced scale that includes a 'no opinion' category.

#### Summary

Scaling techniques can be classified as comparative or non-comparative. Comparative scaling involves a direct comparison of stimulus objects. Comparative scales include paired comparisons, rank order, constant sum and the Q-sort. The data obtained by these procedures have only ordinal properties. Verbal protocols, where the participant is instructed to think out loud, can be used for measuring cognitive responses. In non comparative scaling, each object is scaled independently of the other objects in the stimulus set. The resulting data are generally assumed to be interval or ratio scaled. Non-comparative rating scales can be either continuous or itemized. The itemized rating scales are further classified as Likert, semantic differential or Stapel scales. The data from all these three types of scale are typically treated as interval. Thus, these scales possess the characteristics of description, order



and distance. When using non comparative itemized rating scales, the researcher must decide on the number of scale categories, balanced versus unbalanced scales, an odd or even number of categories, forced versus non-forced choices, the nature and degree of verbal description and the physical form or configuration. Multi-item scales consist of a number of rating scale items.

 $\%\,\%\,\%\,\%\,\%\,\%\,\%\,\%\,\%\,\%\,\%\,\%\,\%$ 

