

SCALING TECHNIQUES IN SOCIO-LEGAL RESEARCH

*B.N. Ghosh**

Introduction

IN SOCIO-LEGAL research, it is sometimes very essential to make distinctions of degree rather than of quality. The technique for registering difference in degree may be of two types : Firstly, one can make a judgment about some characteristic (say criminality) of an individual and place him on a scale which is meant for measuring that characteristic. A scale is a continuum from highest to lowest points, and has intermediate points in between these two extremes. The scale points are so related that the first point indicates a higher position than the second. The second point is higher than the third point and so on. Secondly, the technique may consist of questionnaires prepared in such a fashion that the score of individual's responses gives him a particular place on the scale.

A scale is a method of measurement. In socio-legal research, attitude, behaviour and other qualitative characteristics can be measured by means of different scales. The scalability of a phenomenon depends on its continuum. The various factors of the phenomenon must be logically interrelated and should be capable of continued measurement. A scale must be reliable.

Types of scales

Measurement scales

There are four main types of measurement scales, viz., nominal, ordinal, interval (cardinal) and ratio, which are discussed below :

Nominal scale

A nominal scale consists of two or more named categories into which individuals, objects or responses are classified. In a nominal scale, it is possible to distinguish two or more categories relating to the

* M.A., Ph.D., Department of Economics, Panjab University, Chandigarh.

specified attribute. The members of these categories differ with respect to the specified attribute which is being measured. It is a simple method of classification rather than an arrangement along a continuum. The question of dimension is not important in this type of scale. If desired, the different groups may be numbered. Classification of individuals according to religion is an example of nominal scale.

Ordinal scale

In this type of scale, numbers, i.e., 1,2,3... are assigned to indicate only the relative position. The scale purports to give ranks to the individuals along the specified continuum. But such a scale, like the nominal scale, does not measure the distance between the positions. In this scale, it is very essential to determine the order of position (in terms of more or less, better or worse and so on) in relation to the attribute which is being measured. For instance, X is regarded as more rational than Y. The ordinal scale can indicate only a person's relative position, but it cannot cardinally measure the differences between persons. One can say, for example, X is greater than Y, but he cannot say by how much.

Interval (cardinal) scale

This scale has equal units of measurement and it is possible to interpret not only the order of scale scores but also the distance between them. Thus, two persons with scale positions 4 and 5 are as far apart as two persons with scale positions 10 and 11. A person having the score of 10 cannot be regarded as two times higher or better than a person whose score is 5. When we talk of 10°C rise in temperature, we should not mean that it is twice as hot as 5°C temperature. In an interval scale, the intervals remain equal.

Ratio scale

This scale incorporates the properties of an interval scale together with a fixed origin or zero point. Weights, lengths and times are obvious example. On the basis of a ratio scale, one can compare both differences in scores and the relative magnitude of scores. For instance, the difference between ten and fifteen minutes is the same as that between fifteen and twenty minutes, and ten minutes is twice as long as five minutes.¹

1. Claire Selltiz, Marie Jahoda, Morton Deutsch and Stuart W. Cook, *Research Methods in Social Relations* 186-95 (Revised one-volume ed., Holt, Rinehart and Winston, Inc. New York, 1964).

Rating scales

The rating, ranking and attitude scales have one common method of assigning numerical positions to individuals so that variations in degree may be ascertained. While preparing a rating scale, the rater places the individual at a particular point along a continuum, and a numerical value is attached to the point. The following are the main types of rating scales which may be used in socio-legal research.

Graphic rating scale

Under this method, the rater indicates his rating by putting a tick at the point selected by him on a line chosen for measuring an attribute and specifying points from lowest to highest. The different scale points may indicate brief descriptions about the attitude levels. This may help the rater indicate his own preference. The following is an example of a graphic rating scale :

Prisoners' participation in jail management

1	2	3	4	5
Feels strong for necessity prisoners' management participation	Generally feels the necessity of participation	Partly favours, partly disfavours	Generally feels no necessity of participation	Strongly no of

The rater may mark any one of the five points to indicate his own rating and thereby may reveal his preference. The graphic rating scale is a very simple and convenient device to use. It ensures a fair level of fineness of scoring. In this scaling, vague, ambiguous and practically unlikely and extreme statements should be avoided. Secondly, as far as possible, the descriptive statements should closely correspond to the numerical points on the scale.

Itemised rating scale

This is also known as the "specific category scale" or "numerical scale". In this scale, the rater has to select one of the limited number of categories that are in order in terms of their scale position. Generally, in such a scale, there are five or seven categories. As a general rule, the more clearly defined the categories, the more reliable become the ratings. The number of specifications, however, depends on the nature of the research problems.

Comparative rating scale

In the case of comparative rating scale, the positions on the rating scale are explicitly defined in terms of a given population or group, or in terms of people with known characteristics. Thus, the rater may be asked to specify the comparative ability of a judge with reference to the judgment in a court or he may be asked to specify the comparative skill of an investigator with reference to other investigators. The comparative ability of the individual or the group in question may be expressed in terms of percentage by the rater. For instance, A is more competent than 10 per cent of the judges. The rater must have a clear knowledge of the abilities of the given groups or individuals.

The raters themselves must be unbiased and trained. The specifications of the reference groups and the definitions of the attributes being measured must be objective and clear. In the case of judgment of complex attributes, it is better to use a less differentiated scale. Several raters giving independent judgment can work as a team and can be very helpful in increasing the reliability of ratings, particularly when these are made during the process of analysis by coders. Be that as it may, the rating scales are very simple and useful to apply irrespective of the method of data collection.²

Rank order scales

Rank order scale is another method of comparative and relative rating. In this method, the judge is required to rank individuals in relation to one another. When the population is very limited the judge has to prepare a rank order of individuals from highest in the scale to the lowest.

In a rating scale, the individual rater himself may be the subject of rating. This is known as self-rating which has its own advantages as well as limitations. However, self-rating has been found to be useful in measuring the attitudes, such as intensity, importance and liking. The following two methods may be illustrative.

Paired comparison

This is a simple method of ranking scale. In this type of scale, two stimuli are presented before the judges, out of which the other one is to be selected. The continuum is properly defined. The ranks for different types of jobs, for instance, suitable for recidivists, can be determined. The investigator can make several pairs of jobs available. The respondent may be asked to point out which of the two jobs he/she likes. After all the pairs have been considered, any possible inconsistency may be

2. *Id.* at 345-50.

located. For instance, if job 2 is preferred to job 3, and job 3 is preferred to job 4, then by transitivity axiom, job 2 is preferred to job 4. If somebody prefers in this connection job 4 to job 3, his behaviour is inconsistent and the inconsistent result may be eliminated. In this method, the number of preferences of all persons for a particular job is added and is divided by the number of people who are giving the preferences. The result that is obtained becomes the scale value for that particular job. On this basis, score values are ascertained. The scale value is explained numerically as follows.

<i>Persons preferring</i>	<i>Number of preferences</i>
A	6
B	5
C	5
D	4
E	3
F	1

The scale value= $24/6=4$

This method is only a rough and simple method. It can, however, be made more meaningful by constructing a paired comparison matrix, as has been done by J.P. Guilford in his *Psychometric Method*.³

Horowitz method

Horowitz applied a ranking scale for testing the racial prejudices. He took 8 pictures of Negroes and 4 pictures of white children. These 12 pictures were shown to the school children who were asked to indicate their preferences. First preference was indicated by number 1, second by 2, third by 3, and so on. Then, separately, the scores for the white and for the Negroes were added and compared. In this way, it was possible to know the attitudes towards Negroes and Whites. Needless to say, the smaller the score, the greater is the preference. If all the White children are given the first four preferences, the total score would be 10 ($=1+2+3+4$) but if they are given the last four preferences, the total score would be 42 ($=9+10+11+12$). So the possible range for the White children would be from 10 to 42. The probability of score value would, therefore, be $\frac{10+42}{2} = 26$. If the actual score is less than 26, it would

3. (McGraw-Hill Book Co., Inc., New York, 1954).

be considered as a favourable attitude, and a score more than 26 would imply an unfavourable attitude.

The same photos can be used for choosing the companions for a variety of situations, case by case. If there are only 5 situations, the maximum number of White faces in all the five situations (say, for going to party, for play, for dancing) would be $5 \times 4=20$. Now, taking 20 as equal to 100, the percentage of actual number can be calculated. Since the number of White faces was 4 out of 5, the probability of selection was 80 per cent. Any increase or decrease in the number may be in favour or against the case depending on the direction of difference. Picture tests have been used by many investigators to find out the decisive factors governing individual preferences and attitudes.

Attitude scales

In this type of scale, the attitude of an individual towards a matter, thing, object or system can be known from the score of his responses given on a questionnaire. The score will place him on a scale. He simply expresses his liking or disliking, agreement or disagreement with the issue involved, as given in the forms of questions. On the basis of this reply, he is assigned a score which indicates his position. This type of technique is used in measuring the social attitudes.

In attitude scale, some relevant statements are to be considered by the respondents. The statements are formed in such a way that they are intimately related to the attitude which is sought to be measured. Indirect statements may sometimes be used to reveal the attitude. Secondly, the scale should be able to specify the various but crucial shades of opinions. For instance, extreme, moderate and neutral attitudes must be expressed through the statements so that the respondents have several reasonable ways of expressing their attitudes. Thus, the method of scale formulation remains very decisive, and so is the method of scoring. Depending on the differences in these regards, there may be various types of attitude scales. Let us discuss them.

Point scale

In this scale, at first, a crucial number of words about which the opinion is required is selected. The words may be dowry, eve-teasing, prostitution, divorce and so on. The respondent is to cross out every word that is more annoying than pleasing to him. One point is given to each agreement or disagreement whichever is to be chosen. If agreement is to be considered, for instance, each word which has not been crossed will be given one point. According to this method, the attribute of a respondent is known by calculating the number of words crossed or not

crossed. The words selected should be suggestive of an attitude and the opposite words should also be given at the same time.

There may be another variety of point scale. In such a scale, two sets of words, indicating both favourable and unfavourable opinions, are given. The unfavourable items may be crossed and favourable items may be left uncrossed. For instance, if a person scores out words favouring capitalism and leaves out those favouring communism, he can be said to be a communist.

Differential (Thurstone) scale

This scale is associated with the name of L.L. Thurstone. Hence, this is also known as Thurstone scale. In this scale, a number of statements whose position on the scale has been determined by judges, is used. The position is determined by the method of equal-appearing intervals. The procedure is as follows.

At first, a large number of statements relating to the attitudes are collected by the researcher. A large number of judges work independently to classify these statements into eleven groups. In the first group, all the statements which are most unfavourable to the specified issue, are placed. The next unfavourable statements are placed in the second group, and so on. The statements of the eleventh group are considered most favourable. The sixth position on this continuum is the point at which the attitude is neutral. The first group is given score 11, and the eleventh group is given score 1. The scale value of a statement is computed as the mean or median position to which it is assigned by the judges. Statements which are ambiguous, vague, irrelevant and over which judges differ widely, are discarded. Finally, the scale is prepared by taking into account the evaluated statements that spread out evenly from one extreme to the other.⁴

Thurstone scale thus consists of a series of statements whose positions have been determined neutrally by the judges. At the time of administration of scale questionnaire, the respondents are asked to check the statement or statements with which they agree. The scale values are not shown in the questionnaire and the statements are arranged randomly. The mean or the median of the scale values of the items which are checked by the respondent indicates his position in the scale. This type of scale has been widely used to measure attitudes towards various social phenomena, e.g., dowry, prostitution. The scattered responses of an individual imply that the respondent has no definite and organised attitude towards the phenomenon. Thurstone scale is, however, most

4. L.L. Thurstone, "Attitudes Can Be Measured", *American Journal of Sociology* 529-54 (1928).

appropriate and reliable in the case of a single attitude which is sought to be measured.

Summated (Likert) scale

This type of scale frequently used in the measurement of social attitude was first devised by Likert. Hence, it is also known as Likert-type scale. Unlike the differential scale, the Likert scale uses only the definitely favourable and unfavourable statements. Likert scale excludes the intermediate opinions. This scale consists of a series of statements to which the respondent is to react. The respondent indicates the degrees of agreement or disagreement. Each response is given a numerical score, and the total score of a respondent is found out by summing up his different scores for different responses. This total score indicates his position on the continuum.

The Likert scale uses several degrees of agreement or disagreement, e.g., strongly approve, approve, undecided, disapprove, strongly disapprove. These five points will constitute the scale. Each point of the scale carries a score "strongly approve" is given the highest score (5 or +2) and "strongly disapprove" is given the least score (1 or -2). Other points will have the scores accordingly (i.e., 5, 4, 3, 2 and 1). The score values are not indicated in the questionnaire. The following is a Likert-type scale.

Prohibition should be made compulsory

1	2	3	4	5
Strongly approve	Approve	Undecided	Disapprove	Strongly disapprove
(5 or +2)	(4 or + 1)	(3 or 0)	(2 or - 1)	(1 or - 2)

The following is the method of construction of a Likert-type scale :

- (i) The researcher gathers a large number of statements which clearly indicate favourable or unfavourable attitude towards the issue in question.
- (ii) The questionnaires consisting of the above five points with respect to a statement are administered to the respondents who indicate their responses.
- (iii) The responses will imply various scores as shown above. The scores are consistently arranged either from the highest to the lowest, or from the lowest to the highest.

- (iv) By adding up the different scores of an individual, his total score is calculated, *i.e.*, summation of different scores for different statements.
- (v) The researcher should identify the items which have a high discriminatory power. The responses are interpreted to determine which of the statements discriminate very clearly between high scores and low scores on the total scale. It has to be ensured that the questionnaire is internally consistent. With a view to achieving this, the items with low discriminatory power or those having no significant correlation with total score are eliminated. The main idea is to ensure that every item or statement is related to the attitude which is under study.⁵

Cumulative (Bogardus scale)

In a cumulative scale, a respondent is shown a number of questions to express his agreement or disagreement over an issue. The items are arranged in such a way that a respondent who responds favourably to item number 2 also replies favourably to item number 1 and one who replies favourably to item 3 also replies favourably to items 1 and 2, and so on. Therefore, the individuals who answer favourably have higher total score than those who answer unfavourably. The score of an individual is computed by counting the number of items he answers favourably. His scores indicate for him a particular position on the scale. The intervals between the positions may not be equal. The items may be arranged from favourableness to unfavourableness in a systematic manner or may be randomly selected.

The cumulative type of scale was successfully used by Bogardus first. Thus, it is also known as Bogardus social distance scale. The main purpose of social distance scale is to measure the attitude towards a particular racial group or groups. A number of suggested relationships may be listed, to which members of an ethnic group may be admitted. The respondent is to indicate as to which racial group is to be admitted by him for each of the specified relationships. The attitude is measured by the closeness of relationship that a respondent is willing to accept or the social distance that he likes to maintain. The Bogardus-type scale is given below :

5. See R.L. Wilder, *Introduction to the Foundations of Mathematics* 45-49 (Wiley, New York, 1952).

<i>Relationship I would accept</i>	<i>Hindus</i>	<i>Sikhs</i>	<i>Muslims</i>	<i>Christians</i>
1. Marriage	1	1	1	1
2. Friendship	2	2	2	2
3. Neighbour	3	3	3	3
4. Employment in same occupation	4	4	4	4
5. Membership of club	5	5	5	5
6. Guest to my club	6	6	6	6
7. Exclusion from my club	7	7	7	7

The respondent is to circle each of the seven categories to which he is willing to accept a particular group. The respondent's first feeling reactions can be known by this. For a group, if a respondent circles 3, he is also expected to circle 4 and 5 for the same group. If a respondent does not circle 3, he will most probably not circle 1 and 2, for these indicate even closer relationship (for the same group). The seven categories indicate a gradually increasing social distance.

The fact that the specified items in Bogardus scale form a cumulative scale is empirically borne out in many countries. However, some practical reversals may be noticed on the social distance scale. For instance, a group may not be liked as neighbour, but may be liked as friend in a club. But these reversals are more possible in individual cases rather than for the group as a whole. The reversals have been found to be mostly due to the influences of external factors.

In the Bogardus-type scale, the respondent has to indicate his first feeling. He has to give his reaction to each race or religion as a group and he should not take into account any individual member of a group, best or worst, into account. The social distance can also be calculated mathematically. In order to do this, weights are to be attached to different categories of relationships. Thus, if there are only five categories, the weights such as 1,2,3,4 and 5 can be assigned to the first five categories respectively. The following procedure is generally adopted for the measurement of social distance :

- (i) Place the weights and percentage response for each category in rows;
- (ii) multiply the percentage response by its weight; and
- (iii) add up the product; and this will be the social distance (see the following chart).

Cate-Weight gory		Hindus		Sikhs		Muslims		Christians	
		% Resp- onse	Weights x %	% Resp- onse	Weights x %	% Resp- onse	Weights x %	% Resp- onse	Weights x %
1	1	90	90	45	45	10	10	5	5
2	2	94	188	50	100	20	40	10	20
3	3	95	285	60	180	30	90	15	45
4	4	96	384	70	280	40	160	20	80
5	5	97	485	80	400	50	250	25	125
Total			1432		1005		550		275

In Bogardus-type scale, the score does not indicate the exact extent or degree of preference of a group over the other. This is, of course, the implicit idea of any ordinal scale.⁶

With the development of Thurstone and Likert scaling methods, attention has been shifted away from cumulative scaling towards unidimensional scaling which seeks to avoid the influence of the external factors. In fact, doubts have been raised by Carter and others regarding the basis of continuum scale as given by Thurstone-type scale. Different statements, for example, regarding war, in a Thurstone-type scale do not fall along a straight line. For instance, the statements—"war is good", "war is bad" and "I do not think about war"—cannot make one continuum or a straight line scale. When one is combining different and contradictory statements about an issue, it is impossible to specify as to what is exactly being measured. This reaction has given rise to the development of various scaling techniques and Guttman scale is one such attempt in this direction.

Scalogram (Guttman method)

The Guttman scale is based on the assumption that the various attitude statements in the scale belong to the same dimension. The attainment of a high degree of unidimensionality is the major concern of Guttman scale. However, Guttman scale belongs to the broad category of cumulative scaling. According to Guttman, a universe of content can be considered to be unidimensional only if it yields a perfect or nearly perfect, cumulative scale, that is, whether it is possible to arrange the responses into a pattern of the following type :

6. See B. Phillips, *Social Research* 223-25 (Macmillan, New York, 1971).

Scalogram

Score	Says "Yes" to item			Says "No" to item		
	3	2	1	3	2	1
3	x	x	x			
2		x	x	x		
1			x	x	x	
0				x	x	x

If the above pattern holds, then a given score on a particular series of items always has the same meaning. A score enables one to say as to which item is endorsed by the respondent. For instance, a man having a score of 2 will say yes to statements 2 and 1, but will say no to statement number 3. In this way, some uniformity may be achieved. The scalogram technique is based on reproducibility criterion, *i.e.*, it is possible to reproduce the responses of the respondent about each item from the score item. This is a major test of Guttman scale (see the chart above).

Perfectly cumulative or unidimensional scale is hardly possible in socio-legal research. The scalogram analysis is a simple method of testing the scaleability of the statements. The views of judges are not necessary in this case. A diagram in which individual responses are laid out is called a scalogram. In the Guttman technique, the perfect scale implies that a person who answers a given question favourably will have a higher total score than a person who answers it unfavourably. Guttman scale is analytically complex, apart from the fact that there is no guarantee that the various items will scale, and even if they do, the universe of content may remain narrow in coverage. Guttman method is more appropriate for scaling ordered behaviour than less structured and broad based attitudes.

Guttman model is deterministic in nature. It assures that a person who responds positively/negatively to one item, must respond positively/negatively to a series of others. The model can be made probabilistic rather than deterministic. This attempt has been made by "Latent Structure Analysis", as developed by Lazarsfeld.

Guttman has developed another technique. According to this, the respondent not only gives his view (agreement or disagreement) but he also mentions the intensity (degree) which is classified into five categories. For each respondent, we thus get two scores—content score and intensity score—which can be plotted against each other on a group. The result often gives a U-shaped curve. The more extreme views have highest intensity. The content score at the lowest point of the curve may

be regarded as the dividing line between favourable and unfavourable responses. The lowest point suggesting favourable response is not afflicted by the form and the wording of the individual items. Hence, this method is objective.⁷

Recently, two more scaling techniques—the Q-sort and Semantic differential scaling—have been introduced in the area of attitude scaling.⁸

Conclusion

Attitude scaling is essentially a very vast area of research which is in constant flux. We have only outlined very simply some of the existing methods. A researcher should first try to employ one of the existing scales in his enquiry. But the scale must be appropriate to his population and the subject he is studying. This may save his time and energy. But given ingenuity, a researcher may develop useful measuring technique and contribute to the existing stock of knowledge in this area.

7. L. Guttman, "Principal Component of Scale Analysis", in Stouffer *et al.* (Eds.), *Measurement and Prediction* 312-61 (Princeton, 1950).

8. See *supra* note 6 at 212-14.