CHAPTER - IV

METHODOLOGY

Based on the hypothesis and the literature presented earlier, it is Predicted that a celebrity source will have a greater impact than a non-celebrity endorser on the responses to the advertisement and the advertised product, it is expected that the celebrity will be seen as more credible and that the message will be rated more favourably along the evaluation dimensions. The positive response is expected to produce a more favourable attitude toward the product and eventually a greater intention to use it.

The experimental manipulation varied the ad copies featuring celebrity and non-celebrity model endorsing a particular brand of soft drink, where two versions of nearly identical pair of ads featuring a celebrity and a non-celebrity model, non of which has ever so far appeared in any campaign in the Indian market and no one of the subjects reported seeing any of the messages before the study was conducted. In the first variation (ad copy) an internationally known pop singer was shown endorsing an existing soft drink in a 30 second slot T.V commercial, while the other ad copy was showing an unknown/anonymous model endorsing the same brand of soft drink.

Subjects were asked to complete a questionnaire to indicate their innitial beliefs and attitudes followed by repeated number of exposures to test the ad through the local T.V channel (Video Cable) during the day and night feature

film times for a period of 30 days, which was again followed by a questionnaire administration to indicate the post treatment effect on the status of attitudes of the subject audience.

The sample was composed of 197 subjects who were assigned into the celebrity and non-celibrity treatment groups, demographically, the sample was fairly typical in terms of sex, education, occupational status, age, and geographical dispersion was attained by selecting the samples from completly different ereas in Baroda City- Gujarat state.

The Experimental Design of the study was ran as 3 groups 6 observation (3 endorsement types, pre and past measure for each group) with 81 subjects per each experimental group and 35 for the control group resulting in a total sample size of 197 respondents. The investigator approached each of the 197 subjects and asked them to indicate their attitude toward the advertisement by answering the 40 statment Likert type questionnaire.

THE QUESTIONNAIRE DESIGN :

To measure the attitudes of a large number of individuals it is necessary to construct a special measurement scale. Such as a questionnaire that constitute statements regarding an object in target for the study. It is assumed that the research can classify the statements into

two classes, favourable and unfavourable, with approximately the same number of statements in each class. These statments were then given to a group of subjects who were asked to respond to each one in terms of their own agreement or disagreement with the statment. In obtaining the responses from the subjects the researcher permitted them to use any one of five categories : Strongly agree, agree, uncertain, disagree, strongly disagree. For any given statment the researcher has available proportion of subjects given each of the five categories of response. Thus it was required to weight these categories of response in such a way that the response made by individuals with the most favourable attitudes will always have the highest positive weight. For the favourable statements, it was assumed that this is the "strongly agree" category, and for the unfavourable statements, it is assumed that it is the "strongly disagree" category. Table 4.1 illustrates the process in which the proportion of subjects falling in each response category for the favourable statements is shown in row (1). In row (2) the table cumulative proportions are given, and in row (3) the proportion below a given category plus 1/2 the proportion within the category. For example the second entry in row 3 is obtained by 0.0+1/2(.06)=.03 from the table of the normal curve it is possible to find the deviates to the proportions of row (3).

Table 4.1 The distribution of subjects falling in different response categories. (for a favourable statement)

	Strongely Disagree	Agree	Uncer-tain	Agree	Strongely Agree
1) P	0.0	.06	.22	.18	.04
2) PC	0.0	.06	.28	.46	.50
3) midpoint PC	0.0	.03	.17	.37	.48
4) Z	-3.090	-1.881	-.954	-.332	.050
5) Z+3.090	0.0	1.209	2.136	2.758	3.04
6) Z round-ed	0	1	2	3	3

The proportion of subjects (N=50) falling in each of five response categories for a favourable statment and the normal deviate wieghts for these response categories based upon the proportions.

The normal deviates are shown in row (4) and they are one set of weights the researcher might use for the response categaries. The researcher can make the weights all positive by adding the absolute value of the longest negative value, 3.090, to all of the other entries in row (4) thus obtaining the values shown in row (5). it will be observed that if one rounds the entries in row (5) to the nearest integer, the weights obtained will be 0,1,2,2,3 & these are close to the values 0,1,2,3,4.

Table 4.2 will illustrate that the same thing will happen when one deals with the responses to an unfavourable statement, although here the researcher has reversed the weightings for the response categories so that the strongly disagree category has the highest positive weight.

Table 4.2 The Distribution of respondents falling in different response categories. (for an unfavourable statement)

	Strongely Agree	Agree	Uncertain	Disagree	Strongely Disagree
1) P	.2	.30	.06	.14	.06
2) PC	.2	.32	.38	.52	.58
3) midpoint PC	.1	.17	.35	.45	.55
4) Z	-1.282	-.954	-.385	-.126	.126
5) Z+1.282	0.00	0.328	0.897	1.156	1.408
6) Z rounded	0	0	1	1	1

The proportion of subjects (N=50) falling in each of five response categories for an unfavourable statement and the normal deviate weights for these responses categories based upon the proportions.

In the development of the method of attitude * scale described in this study. Likert (1932) states that

*Likerts' (1932) monograph reporting his research was subsequently reprinted (with a few changes in a volume by Murphy and Likert (1937). The latter publication also contains a more detailed report of applications of scales constructed by the likert technique.

the scores based upon the relatively simple assignment of integral weights correlated .99 with the more complicated normal deviate system of weights. He therefore used a simple system which has been adopted in this study. For favourable statements, the strongly agree response will be given a weight of 4, the agree response a weight of 3, the uncertain response a weight of 2, the disagree response a weight of 1 and the strongly disagree response a weight of 0. For unfarourable statements, the scoring system is reversed with the strongly disagree being given the 4 weight and the strongely agree response the 0 weight.

For each subject the researcher obtains the a total score by summating his scores for the individual items. Because each response to a statement may be considered a rating and because these are summated over all statements. Bind called the Likert method of scale construction the method of summated ratings.

As quated in the techniques of attitude scale construction[2]. The researcher has a basis for rejection of statements in terms of Q and the criterion of irrelevance. As basis for rejecting statements in the method of summated ratings, use is made of some form of item analysis. It is considered as the frequency of distribution of scores that is based upon the responses to all statements. Then 25 per cent

of the subjects is taken, out of population of 100 that was selected from different colleges in Baroda City in which 35 students from Faculty of Management Studies M.S. University and 35 students from Faculty of Commerce post graduate students while the rest were selected from a management school - with the highest total scores and also the 25 per cent of the subjects with the lowest total scores, these two groups were treated separately for the assumption that these groups will provide criterion groups in terms of which to evaluate the individual statements. In evaluating the responses of the high and low groups to the individual statements the ratio was to be found as under :

$$t = \frac{\bar{X}_H - \bar{X}_L}{\sqrt{\frac{S'^2_H}{n_H} + \frac{S'^2_L}{n_L}}}$$

where $\bar{X}_H$ = the mean score on a given statement for the high group.

$\bar{X}_L$ = the mean score on the same statement for the low group.

S^2_H = the variance of the distribution of responses of high group to the statement.

S^2_L = the variance of the distribution of responses of low group to the statement.

n_H = the number of subjects in the high group.

n_L = the number of subjects in the low group.

if $n_H = n_L = n$, as would be the case if we select the same percentage of the total number of subjects for the high and low groups, then the following formula can be written :

$$t = \frac{\bar{X}_H - \bar{X}_L}{\sqrt{\frac{\sum(X_H - \bar{X}_H)^2 + \sum(X_L - \bar{X}_L)^2}{n\,(n-1)}}}$$

where $\sum(X_H - \bar{X}_H)^2 = \sum X_H^2 - \frac{(\sum X_H)^2}{n}$

and $\sum(X_L - \bar{X}_L)^2 = \sum X_L^2 - \frac{(\sum X_L)^2}{n}$

This equation will illustrate the calculation of t for evaluating the difference in the mean response to an attitude statement by a high group and a low group[3].

	Low Group				
Response category	X	X^2	f	fx	$f(x)^2$
Strongely Agree	4	16	1	44	16
Agree	3	9	15	45	405
Uncertain	2	4	6	12	48
Disagree	1	1	2	2	2
Strongely Disagree	0	0	1	0	0
sums			25	63	471
			n_L	$\sum X_L$	$\sum X_L^2$

High Group

Response category	X	X^2	f	fx	$f(x)^2$
Strongely Agree	4	16	0	0	0
Agree	3	9	3	9	81
Uncertain	2	4	11	22	88
Disagree	1	1	9	9	9
Strongely Disagree	0	0	2	0	0
sums			25	40	178
			n_H	ΣX_H	ΣX^2_H

The calculation of t for evaluationg the difference in the mean response to an attitude statement by a high group and a low group.

$$\bar{X}_L = \frac{63}{25} = 2.52 \qquad \bar{X}_H = \frac{40}{25} = 1.6$$

$$\Sigma(X_L - \bar{X}_L)^2 = \frac{471 - (63)^2}{25} = 312.24$$

$$\Sigma(X_H - \bar{X}_H)^2 = 178 - \frac{(40)^2}{25} = 114$$

$$t = \frac{1.6 - 2.52}{\sqrt{\frac{114 + 312.24}{25\,(25 - 1)}}} = -.129$$

The value of t is a measure of the extent to which a given statement differentiates between high and low groups. It was regarded that any t vlaue equal to or greater than 1.75 showed an indication that the average response of the high and low groups to a statement differs significantly, provided that 25 (out of 100 respondents) or more subjects were considered in the high and low groups.

In this method (method of summated-ratings) it was desired to have a set of 20 to 25 statements that would differentiate between high and low groups. These statements were selected by finding the t value for each statement and then arranging the statements in rank order according to their t value. Then select the 20 to 25 statements with the largest t value for the desired likert type attitude scale.

Other alternative methods of item analysis

Other methods of item analysis, such as correlated method could have been used in evaluating the individual statements instead of the t test described earlier. However it is stated that it is doubtful whether any of item analysis in current use would result in an ordering of the statments that is essentially different from the ordering detailed in terms of t values, Indeed, often a simple procedure than t test will prove to be sufficient.

According to Murphy and Likert, the rank ordering of 15 statement upon the basis of the magnitude of the difference between the means of a high and low group agreed very well

with the magnitude of the correlation between the item response and total score. As a simple and convenient procedure, therefore, the difference between the means of the high and low groups on the individual statements might be used as the basis for selecting the 20 to 25 items desired for the scale.

It is recommended that approximately half of the selected statements should be favourable so that the strongly agree response carries the 4 weight and the strongely disagree response the 0 weight. The other half should consist of unfavourable statements so that the scoring system is reversed.

The advantage of having both kinds of statements represented in the final scale is to minimize possible response sets of subjects that might be generated if only favourable or unfavourable were included in the scale.

The scale then was distributed to a new group of subjects who constituted the real experimental sample to be studied before and after the treatment.

The reliability of the scores on the scale was obtained by correlating scores of the favourable (odd numbered) statements with those of unfavourable statements.

The reliability coefficients typically reported for scales constructed by the method of summated-ratings are above .85 , even when fewer than 20 items make up scale.

INTERPRETATION OF SCORES

The interpretation of an attitude score in the scale cannot be made independently of the distribution of scores of some defined group. If a subject obtains a score of 0 on a 25 - item summated scale, it interpretes the score as indicating an unfavourable attitude, since, in order to obtain this score, the subject would have had to have given a strongly agree response to every unfaourable statement and a strongly disagree response to every favourable statement in the scale. Similarly, it could be interpreted that a scale that a score of 100 as indicating a favourable attitude, since this score could be obtained only if the subject gave a strongly agree response to every unfavourable statement and a strongly disagree response to every favourable statement in the scale. The interpretation of the score falling between the maximum and minimum possible scores is more difficult, if the concern was to describe an individual as having either a favourable or an unfavourable attitude toward the object under consideration. That is because the summated-rating score corresponding to the zero or "neutral" point on a favourable-unfavourable continum is not known as it is assumed to be known in the case of equal intervl scores. Nor is there any evidence to indicate that "neutral" point or a summated-rating scale necessarily range of scores, that is, to the score of 50 on a 25 item scale.

It is important to mention that the absence of knowledge of such a point is a handicap only if the major

interest is in being able to assign, on the basis of an attitude score, a single subject to the class of those favourable objects under consideration.

If the research interest, as is the case in this study, to compare the mean change in attitude score as a result of introducing some experimental variable, such as a motion picture film - T.V ad commerical - then the lack of a zero point should cause no concern. Similarly if the interest was to compare the mean attitude scores of two or more groups, this can be done with summated rating scales as well as with equal-appearing interval scales.

RELIABILITY OF ATTITUDE SCORES

The scale constructed by the method of summated ratings and by the method of equal appearing intervals, largely as a result of Likert's study of the reliability of a Thurstone type scale which was scored by both methods, have had some confusion centered around the subject of comparative reliability.

Two forms of the scale were given to a group of subjects with instructions to check the statements in accordance with the usual Thurstone instructions. The same scales were then given to the subjects with instruction to check for each item one of the five alternatives (Strongly agree, Agree, Uncertain, Disagree, Strongly disagree) in accordance with the usual Likert type responses and were

omitted when the subjects were asked to check reactions according to the method of summated rating scoring system.

The reliability coefficient between the two forms of the scale (22 versus 22 items), when scored by equal appearing interval method, was .88, corrected by Spearmen Brown formula. The reliability coefficient for the two forms (18 versus 18 items) as scored by the method of summated ratings was .94, corrected by the Spearman-Brown formula. This demonstrates that it is possible to take a scale constructed by the method of equal-appearing intervals and apply to most of the statements the scoring system of the method of summated ratings. When this is done, a somewhat higher reliability coefficient will in general, be obtained Ferguson (1941), however in critizing the method of summated ratings, seems to believe that Likert, because he found a higher reliability coefficient with his method of scoring rather than with the equal-appearing intervals method of scoring, erroneously concluded that "his technique is the better one" (P.52). The higher reliability coefficient obtained by the Likert method of scoring, Forguson notes, may be due to the fact that increasing the number of steps on a psychological scale increases reliability" (P.52). As a matter of record, this is precisely the same expla nation offered originally by Murphy and likert. For the higher reliability coefficient obtained by the Likert method of scoring. The discussion, pro and con, on this point has little bearing upon the question of whether the method of

summated ratings or the method of equal-appearing intervals will yield scales of higher reliability. The real problem concerns the reliability of scales constructed by the two methods not the reliability of a particular scoring scheme isolated from the technique of scale construction of which it is a part.

Ferguson (1939) has quoted Thurstone as reporting the reliabilities of scales constructed by the method of equal-appearing intervals. Under his direction, as being "all over .8, most of them being .9" (P. 670). Ferguson adds that in his own studies he has found reliabilities for equal appearing interval scales ranging from ".52 to .80 for the 20 -item forms and .68 to 89 for the 40 -item forms" (P.670). If we take these coefficients as representative, how do they compare with those reported for scales constructed by method of summated-ratings.

Murphy and Likert (1937, P.48) found reliability coefficient for their internationalism scale of 24 items ranging from .81 to .90. Their Imperlialism scale of 12 statements gave coefficients ranging from .80 to .92 ; the Negro scale of 14 statements yielding coefficients ranging from .79 to .91 . Rundquist and Sletto (1936 ,P. 110) report coefficients ranging from 78 to 88 various summmated-rating scale of 22 statements each.

The Likert-type scales with even fewer statements will give high reliability coefficient is indicated by all

(1934,P. 19) reliability coefficient for his scale of 10 statments measuring altitude toward religion ranged from .91 to.93 ,for his scale of 7 statements measuring attitude toward employers the coefficient ranged from .77 to .87 and his moral scale of 5 statements gave coefficient from .69 to .84.

All these coefficents compare favourably with those obtained from scale constructed by the method of equal-appearing intervals. According to the evidence at hand , there is no reason to doubt that the scales constructed by the method of summated-ratings will yield reliability coefficients as high as or higher than those obtained with scales constructed by the method of equal-appearing intervals.

LIKERT SCALE AND THURSTONE SCALE VALUES OF STATEMENT

From the description of the method of summated-ratings it is clear that no consideration is given to the problem of the scale values of the individual statements,thus it becomes sufficient if the statements relating to a given psychological object can be classified as favourable or unfavourable. The classification determines the direction of the weighting system to be assigned to the response categories. In the scaling methods such as the method of equal-appearing intervals, however, a judging group is required on order to determine first the scale values of the statements on the favourable-unfavourable contiuum. In this

study the questionnaire was distributed among 10 teacners from Commerce and Arts faculty at M. S. University of Baroda to provide the researcher with important feedback regarding the structure of the statements and the assigned values to each one. They were also asked to give comments on the complexity or simplicity of the statements, whether the statements are clear and easy to understand, whether they are relevent or irrelevent. The suggestiones were then considered in rearranging some of the statements or omitting the irrelevent ones. The judging group constitute some senior teachers in the field of Managment Psychology and Economics.

The confusion which followed Likerts scoring of the statements in an equal-appearing interval scale by the weighted response method, unfortunately, was not confined to the subject of reliability it spread to involve the questions of whether or not there is a need for a judging group. Ferguson (1941) for example, seems to believe that Likert implied , as result of obtaining a higher rellability coefficient with his method of scoring than with the Customary Thurstone method of scoring , that he had demonstrated that the method of summated-ratings does away entirely with the need for a judging group.It has been claimed by Likert (1932) that the method of summated ratings is simpler and easier to apply in the development of an attitude scale than in the method of equal-appearing intervals. Some support to this claim has been given by others who have used the method of summated ratings. Hall

(1934), states that he used the method of summated ratings. Hall (1934), states that he used the method of summated ratings in his survey of attitudes of employed and unemployed men because of its relative simplicity. Rundquist Sletto (1936) used the method of summated ratings in developing the attitude scale contained in the Minnesoto survey of opinions and they also express their belief that the method" is less laborious than that developed by Thurstone"[4].

References

1. Edward L. Allen "Techniques of attitude scale construction". (Vakils, Feffer and Simons private limited. Baque Building, 9 sprott road Ballard Estate, Bombay 1. 1969) p. 151.

2. Ibid p. 152.

3. Ibid p. 153.

4. Ibid p. 169.