

11. *The Method of Constructing an Attitude Scale*

RENSIS LIKERT

I. THE SELECTION OF STATEMENTS

Each statement should be of such a nature that persons with different points of view, so far as the particular attitude is concerned, will respond to it differentially. Any statement to which persons with markedly different attitudes can respond in the same way is, of course, unsatisfactory.

The results obtained in constructing the present scales demonstrate the value of the following criteria. These criteria were kept in mind in collecting the statements for the original Survey of Opinions.

1. It is essential that all statements be expressions of *desired behavior* and not statements of *fact*. Two persons with decidedly different attitudes may, nevertheless, agree on questions of fact. Consequently, their reaction to a statement of fact is no indication of their attitudes. For example, a person strongly pro-Japanese and a person strongly pro-Chinese might both agree with the following statements:

"The League of Nations has failed in preventing Japan's military occupation of Manchuria."

or

"Japan has been trying to create in Manchuria a monopoly of trade, equivalent to closing the 'open-door' to the trade of other countries."

To agree with them or believe them true is in no way a measure of attitude.

Rice (27, p. 184) has clearly stated the im-

* Excerpted from the Appendix of "A Technique for the Measurement of Attitudes," *Archives of Psychology*, 1932, No. 140, 44-53, with the permission of the author and the Columbia University Press.

portance of recognizing this criterion when in discussing the Thurstone technique he says:

What is the possibility that the acceptance or rejection by a subject of a statement upon the completed scale may represent a rational judgment concerning the truth or falsity of the statement made? It would seem to exist. If so, the validity of the statement as an index of attitude is destroyed or impaired.

In dealing with expressions of desired behavior rather than expressions of fact the statement measures the present attitude of the subject and not some past attitude. The importance of dealing with present rather than past attitudes has been emphasized by Thurstone (38) and Murphy (22, p. 615). A very convenient way of stating a proposition so that it does involve desired behavior is by using the term *should*.

2. The second criterion is the necessity of stating each proposition in *clear, concise, straight-forward statements*. Each statement should be in the simplest possible vocabulary. No statement should involve double negatives or other wording which will make it involved and confusing. Double-barreled statements are most confusing and should always be broken in two. Often an individual wishes to react favorably to one part and unfavorably to the other and when the parts are together he is at a loss to know how to react. Thus in the following illustration a person might well approve one part and disapprove another part:

"In order to preserve peace, the United States should abolish tariffs, enter the League of Nations, and maintain the largest army and navy in the world."

To ask for a single response to this statement makes it meaningless to the subject. This state-

ment should be divided into at least three separate statements.

The simplicity of the vocabulary will, of course, vary with the group upon whom the scale is intended to be used, but it is a desirable precaution to state each proposition in such a way that persons of less understanding than any member of the group for which the test is being constructed will understand and be able to respond to the statements. Above all, regardless of the simplicity or complexity of vocabulary or the naïveté or sophistication of the group, each statement *must avoid every kind of ambiguity*.

3. In general it would seem desirable to have the questions so worded that the modal reaction to some is more toward one end of the attitude continuum and to others more in the middle or toward the other end. In this manner, the modal value of the responses to the different questions in the attitude scale will range fairly well across the attitude continuum. There is no need, however, to have questions whose modal reactions are at either extreme of the continuum.

4. To avoid any space error or any tendency to a stereotyped response it seems desirable to have the different statements so worded that about one-half of them have one end of the attitude continuum corresponding to the *left* or *upper* part of the reaction alternatives and the other half have the same end of the attitude continuum corresponding to the *right* or *lower* part of the reaction alternatives. For example, about one-half the statements in the Internationalism scale have the international extreme corresponding with "Strongly approve" while the other half has it corresponding with "Strongly disapprove." These two kinds of statements ought to be distributed throughout the attitude test in a chance or haphazard manner.

5. If multiple choice statements are used, the different alternatives should involve *only a single attitude variable* and not several.

II. CONSTRUCTING THE SCALE

It is usually desirable to prepare and select more statements than are likely to be finally used, because after trying the statements upon a group, some may be found to be quite unsatisfactory for the intended purpose. For this reason after selecting a good number of statements they should be given to the group or a

part of the group whose attitudes we wish to measure. The sample used should be sufficiently large for statistical purposes.

For purposes of tabulation and scoring, a numerical value must be assigned to each of the possible alternatives. If five alternatives have been used, it is necessary to assign values of from one to five with the three assigned to the undecided position on each statement. The ONE end is assigned to one extreme of the attitude continuum and the FIVE to the other; this should be done consistently for each of the statements which it is expected will be included in the scale. Thus if we arbitrarily consider the "favorable to the Negro" extreme FIVE and the "unfavorable to the Negro" extreme ONE, the alternative responses to the following statements would be assigned the values shown in Table 1.

Some may object to the designation made, saying that the terms "favorable" and "unfavorable" are ambiguous or that the favorable attitude is just opposite to that here considered favorable.

So far as the measurement of the attitude is concerned, it is quite immaterial what the extremes of the attitude continuum are called; the important fact is that persons do differ quantitatively in their attitudes, some being more toward one extreme, some more toward the other. Thus, as Thurstone has pointed out in the use of his scales, it makes no difference whether the zero extreme is assigned to "appreciation of" the church or "depreciation of" the church, the attitude can be measured in either case and the person's reaction to the church expressed.

The split-half reliability should be found by correlating the sum of the odd statements for each individual against the sum of the even statements. Since each statement is answered by each individual, calculations can be reduced by using the sum rather than the average.

An objective check ought then to be applied to see (1) if the numerical values are properly assigned and (2) whether the statements are "differentiating." One possible check is item analysis which calls for calculating the correlation coefficient of each statement with the battery. If a negative correlation coefficient is obtained, it indicates that the numerical values are not properly assigned and that the ONE and FIVE ends should be reversed. If a zero or very low correlation coefficient is obtained, it indicates that the statement fails to measure

TABLE 1

Numerical Value	"How far in our educational system (aside from trade education) should the most intelligent Negroes be allowed to go?"				
1	(a) Grade school.				
2	(b) Junior high school.				
3	(c) High school.				
4	(d) College.				
5	(e) Graduate and professional schools."				
	"In a community where the Negroes outnumber the whites, a Negro who is insolent to a white man should be:				
5	(a) Excused or ignored.				
4	(b) Reprimanded.				
3	(c) Fined and jailed.				
2	(d) Not only fined and jailed, but also given corporal punishment (whipping, etc.).				
1	(e) Lynched."				
	"All Negroes belong in one class and should be treated in about the same way."				
	STRONGLY APPROVE	APPROVE	UNDECIDED	DISAPPROVE	STRONGLY DISAPPROVE
Value	(1)	(2)	(3)	(4)	(5)
	"Where there is segregation, the Negro section should have the same equipment in paving, water, and electric light facilities as are found in the white districts."				
	STRONGLY APPROVE	APPROVE	UNDECIDED	DISAPPROVE	STRONGLY DISAPPROVE
Value	(5)	(4)	(3)	(2)	(1)

that which the rest of the statements measure. Such statements will be called undifferentiating. Thurstone (38) refers to them as irrelevant or ambiguous. By "undifferentiating" we merely mean that the statement does not measure what the battery measures and hence to include it contributes nothing to the scale. A statement which is undifferentiating for a scale measuring one attitude continuum may be quite satisfactory for a scale measuring another attitude continuum. The following are some of the reasons why a statement may prove undifferentiating:

1. The statement may involve a different issue from the one involved in the rest of the statements, that is, it involves a different attitude continuum.
2. The statement may be responded to in the same way by practically the entire group. For example, the response to the following statement was practically the same upon the part of all students—some two thousand—to

whom it was given: "Should the United States repeal the Japanese Exclusion Act?"

3. The statement may be so expressed that it is misunderstood by members of the group. This may be due to its being poorly stated, phrased in unfamiliar words, or worded in the form of a double-barreled statement.

4. It may be a statement concerning a fact which individuals who fall at different points on the attitude continuum will be equally liable to accept or reject.

It is, of course, desirable in constructing an attitude scale that the experimenter exercise every precaution in the selecting of statements so as to avoid those that are undifferentiating. However, item analysis can be used as an objective check to determine whether the members of a group react differentially to the statement in the same manner that they react differentially to the battery; that is, item analysis indicates whether those persons who fall toward one end of the attitude continuum on

the battery do so on the particular statement and vice versa. Thus item analysis reveals the satisfactoriness of any statement so far as its inclusion in a given attitude scale is concerned.

No matter for what *a priori* reasons the experimenter may consider a statement to belong in a scale, if the statement, when tried on a group, does not measure what the rest of the statements measure, there is no justification for keeping that statement in the battery. After all, we are interested in measuring the attitudes of the members of the group, not those of the experimenter.

There is no reason to expect that the logical analysis of the person who selects the statements will necessarily be supported by the group. Quite often, because of a lack of understanding of the cultural background of the group, the experimenter may find that the statements do not form the clusters or hierarchies that he expected. It is as important psycho-

logically to know what these clusters are as it is to be able to measure them.

The degree of inclusion, i.e. the size of the correlation coefficient between the item and the battery, required for a particular statement will no doubt be a function of the purpose for which the attitudes are being measured. If a general survey type of study is being undertaken the degree of inclusion required will be less than when a more specialized aspect of attitudes is being studied. A similar relationship is to be noted in the measurement of intelligence.

The only difficulty in using item analysis is that the calculation of the necessary coefficients of correlation is quite laborious. The criterion of internal consistency was tried and the results obtained were found to be comparable with the results from item analysis. Table 2 shows a comparison of the results obtained from item analysis and the criterion of internal consistency.

TABLE 2. Comparison of the Results Obtained from the Application of the Criterion of Internal Consistency and Item Analysis to the Negro Scale for Groups "A" and "B" Combined (N = 62)

Column 1	Column 2	Column 3	Column 4	Column 5
1	.69	1.7	2	5
2	.64	1.5	6	6
3	.51	1.7	10	11
4	.18	0.4	14	14
5	.62	1.3	7	8
6	.40	0.7	11	13
7	.12	0.1	15	15
8	.39	1.1	12	10
9	.26	0.9	13	12
10	.65	2.7	5	1
11	.60	1.2	8	9
12	.54	1.4	9	7
13	.67	2.3	4	2
14	.74	2.0	1	3
15	.68	1.6	3	4

ρ (Column 4 vs. Column 5) = +.91

Column 1—Statement numbers.

Column 2—Coefficient of correlation between the score on the individual statement and the average score on all fifteen statements.

Column 3—Difference between the average score of the highest 9 individuals and the lowest 9 individuals.

Column 4—Order of excellence as determined by item analysis based upon the coefficients of correlation shown in Column 2.

Column 5—Order of excellence as determined by the criterion of internal consistency based upon the differences shown in Column 3.

TABLE 3. Criterion of Internal Consistency Applied to the Internationalism Scale for Group "D" (N = 100)

HIGH GROUP	STATEMENT NUMBERS										Five-Point Statements																	
	Three-Point Statements					Five-Point Statements					Five-Point Statements					Five-Point Statements												
Individual	No.	Score	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		
	85	108	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5		
	65	104	4	4	3	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	3	5		
	13	102	4	4	4	4	4	4	4	4	4	4	4	4	5	5	3	5	5	5	5	5	5	5	3	5		
	10	101	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	4	4	5	5	5	5	5	4	5		
	71	101	2	4	4	4	4	4	2	4	4	4	4	4	5	3	4	5	5	5	5	5	5	5	5	5		
	98	100	4	4	4	4	4	4	4	4	4	4	4	4	5	3	5	4	4	5	5	5	5	5	5	3		
	27	98	4	2	4	4	4	4	4	4	4	4	4	4	4	5	3	5	5	5	5	4	4	5	2	5		
	60	98	4	4	4	4	4	4	2	4	4	4	4	4	5	5	3	4	4	5	5	5	5	4	5			
	64	98	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	3	5	5	5	5	5	4	5			
		Sum of 9-high	34	34	35	36	36	36	32	36	32	36	36	36	42	37	35	40	45	45	40	44	45	45	35	41		
		Sum of 9-low	18	20	20	28	24	29	21	20	22	21	34	23	21	24	22	15	31	22	15	22	24	24	17	14		
	D	16	14	15	8	12	7	11	16	10	15	2	13	21	13	13	25	14	23	25	22	22	21	18	27			
	D/9	1.8	1.6	1.7	.9	1.3	.8	1.2	1.8	1.1	1.7	.22	1.4	2.3	1.4	1.4	2.8	1.6	2.6	2.8	2.4	2.3	2.0	3.0	2.2			
	Order	1.5	5	3.5	10	7	11	8	1.5	9	3.5	12	6	6.5	11.5	11.5	2.5	10	4	2.5	5	6.5	9	1	8			
	(3-point statements and 5-point statements treated separately)																											
	LOW GROUP																											
	17	49	2	2	2	2	2	2	4	4	2	2	2	2	1	2	1	1	2	2	1	2	4	4	2	1		
	77	54	2	2	2	2	4	2	2	2	3	4	4	2	3	3	1	4	4	2	1	2	2	2	1	2		
	22	60	2	2	2	2	4	2	2	2	2	4	4	2	4	3	1	3	4	2	2	4	4	2	1	2		
	35	61	2	2	2	4	3	2	2	2	4	4	4	3	2	3	1	3	2	3	1	3	3	2	2	2		
	53	62	2	2	2	4	4	2	2	2	2	4	4	2	3	2	2	2	3	1	3	1	3	1	1	1		
	69	62	2	2	4	2	2	4	2	2	2	4	4	2	3	3	2	2	4	1	1	4	5	2	1	4		
	94	63	2	2	2	4	3	4	2	2	4	4	4	2	2	3	2	2	4	2	2	3	2	2	2	2		
	21	64	2	2	2	4	2	4	3	2	2	4	4	2	2	3	3	2	4	4	2	3	2	2	2	4		
	88	64	2	4	2	4	2	2	2	2	2	4	4	4	4	3	1	4	4	4	2	2	1	2	3			
		Sum of 9-low	18	20	20	28	24	29	21	20	22	21	34	23	21	24	22	15	31	22	15	22	24	24	17	14		

tency. It will be noted that the relation between the order of excellence for the different statements as determined by item analysis and the criterion of internal consistency as expressed by rho is $+ .91$. Since the criterion of internal consistency is much easier to use than item analysis and yet yields essentially the same results, its use is suggested.

In using the criterion of internal consistency the reactions of the group that constitute one extreme in the particular attitude being measured are compared with the reactions of the group that constitute the other extreme. In practice approximately ten per cent from each extreme was used. Table 3 shows the criterion of internal consistency applied to the Internationalism scale for Group D. This criterion acts as an objective check upon the correct assigning of numerical values in that if the numerical values are reversed on a particular statement the extreme high group will score low on that statement and the extreme low group will score high, i.e. we will obtain a negative difference between the two extreme groups on that question. Furthermore, if a statement is undifferentiating it will not differentiate or discriminate the two extreme groups, i.e., the high group will not score appreciably higher than the low group upon that statement.

Finally, on the basis of the results obtained from item analysis or the criterion of internal consistency and having due regard for all the factors concerned, one should select the most differentiating statements for the final form or forms of the attitude test. If, through this selection of the more differentiating statements, statements concerning a particular aspect of the attitude being measured are eliminated, then, obviously, the final scale can only be said to measure the attitude continuum represented by the remaining statements. For example, if it is found by the use of these objective checks that statements concerning the economic status of the Negro involve an attitude continuum other than that of statements having to do with the social equality of the

Negro, the former should not be mixed with the latter. On the contrary, two attitude scales should be constructed. If, on the other hand, these two groups of statements are found to involve the same attitude continuum, they can be combined into a single scale. As previously stated, the degree of inclusion required or desired will generally be a function of the purpose for which the attitude scales are being used.

A sufficient number of statements should be used in each form to obtain the desired reliability. In preparing the final form or forms, it would be desirable to apply the fourth criterion stated under "The Selection of Statements."

Because a series of statements form a unit or cluster when used with one group of subjects which justifies combining the reactions to the different statements into a single score, it does not follow that they will constitute a unit on all other groups of persons with the same or different cultural backgrounds. For example, an examination of the statements in the Imperialism scale will reveal that it contains statements having to do with imperialism both in China and Latin America, and while it is true that these statements form a sufficient cluster to justify their being treated as a unit with the groups used, still with other groups of persons with markedly different attitudes toward China or Latin America it is probable that this single scale would have to be divided into two or more scales.

The ease and simplicity with which attitude scales can be checked for split-half reliability and internal consistency would seem to make it desirable to determine the reliability and examine the internal consistency of each attitude scale for each group upon which it is used. It is certainly reasonable to suppose that just as an intelligence test which has been standardized upon one cultural group is not applicable to another so an attitude scale which has been constructed for one cultural group will hardly be applicable in its existing form to other cultural groups.

BIBLIOGRAPHY

- Murphy, G., and Murphy, L. B. *Experimental Social Psychology*. New York: Harper, 1931.
- Rice, S. A. Statistical studies of social attitudes and public opinion. In Rice, S. A. (Ed.), *Statistics in social studies*. Philadelphia: University of Pennsylvania Press, 1930, 131-192.
- Thurstone, L. L., and Chave, E. J. *The measurement of attitude*. Chicago: University of Chicago Press, 1929.