

# A TECHNIQUE FOR THE MEASUREMENT OF ATTITUDES

BY  
RENSIS LIKERT, Ph.D.  
NEW YORK UNIVERSITY

ARCHIVES OF PSYCHOLOGY

R. S. WOODWORTH, EDITOR

*No. 140*

*Archives of Psychology  
Volume 22 (1932-33)  
pp. 5-55*

NEW YORK

June, 1932

## TABLE OF CONTENTS

	<i>Page</i>
I. Introduction .....	5
II. Problem .....	7
III. Procedure .....	11
IV. Results .....	21
1. The Sigma Method of Scoring.....	21
2. The Simpler Method of Scoring.....	25
3. The Reliability of the Scales.....	28
4. The Validity of the Scales.....	32
5. Comparison of the Simpler Method with the Thurstone Method of Scoring.....	33
V. Psychological Interpretation.....	36
VI. Summary .....	42
Appendix, The Method of Constructing an Attitude Scale .....	44
Bibliography .....	54

# A Technique for the Measurement of Attitudes

## I. INTRODUCTION\*

Attempts to measure the traits of character and personality are nearly as old as techniques for the measurement of intellectual capacity, yet it can scarcely be claimed that they have achieved a similar success. Part, at least, of the difficulty has lain in the statistical difficulties which are encountered when everyday aspects of social behavior, ordinarily handled as qualitative affairs, are treated from the mathematical point of view. The present study, although part of a larger investigation undertaken in 1929 by Gardner Murphy, aims primarily at the solution of a technical problem which has arisen in relation to the quantitative aspects of the study of social attitudes.

The history and present status of research upon personality traits in general, and social attitudes in particular, have been so thoroughly surveyed by Murphy (21, pp. 381-386, and 22, pp. 558-690), Bain (4), Vetter (41), Katz and Allport (16), Watson (43), and others, that no useful purpose would be served in attempting such a study here.

Nevertheless, among the hundreds of efforts to measure social attitudes during the last few years, the careful procedures developed by Thurstone (34, 38) have naturally and rightly received special attention. These are characterized by a special endeavor to equalize the step-intervals from one attitude to the next in the attitude scale, using the familiar methods of psychophysics for such determinations. The Thurstone methods have been shown to yield a satisfactory reliability, and, in terms of correlations between scores and case histories as evaluated by judges, a satisfactory validity (29).

Many obvious affinities appear between the present study and those of Thurstone, yet in a sense the present report

---

\* This study is one of a series under the general direction of Professor Gardner Murphy.

constitutes a radical departure from the concepts which Thurstone has published, as, for example, in the use of judges.

A number of statistical assumptions are made in the application of his attitude scales,—e.g., that the scale values of the statements are independent of the attitude distribution of the readers who sort the statements (38, p. 92),—assumptions which, as Thurstone points out, have not been verified. The method is, moreover, exceedingly laborious. It seems legitimate to inquire whether it actually does its work better than the simpler scales which may be employed, and in the same breath to ask also whether it is not possible to construct equally reliable scales without making unnecessary statistical assumptions. Since so much is being published about attitude measurement, it seems worth while to raise these questions and to report on some results relative to the problem. It is feared that some will mistakenly interpret this article as an "attack" on Thurstone's methods. I therefore wish to emphasize in the strongest terms that I am simply endeavoring to call attention to certain problems of method, and that I am very far from convinced that the present data close the question.

## II. PROBLEM

From an historical point of view one fact deserves emphasis, namely the extraordinary interest in recent years in the problem of the "generality" or "specificity" of character traits—a problem which in its simplest form has to do with the question of the functional independence of social habits. Roughly speaking, the specificity theory would define personality as a composite of many independent habits, while those who emphasize the relative unity of character point to significant general factors which make for effective prediction from behavior in one situation to behavior in a different situation. It seems to the present writer that the problem of specificity has been greatly misunderstood by many writers upon social attitudes and that the value of such contributions as the present may lie chiefly in redirecting attention to those aspects of the specificity-generality problem upon which quantitative results are clear-cut and psychologically significant. It will be well first to consider briefly the logic of the specificity-generality problem, then to present and evaluate our data, and finally to undertake to find the psychological significance of our data in relation to the problem raised.

If an attitude be defined by the social psychologist as a tendency toward a particular response in a particular situation, it is clear that the number of definable attitudes existing in a given person at a given time will depend upon the range of stimuli to which he is subjected. But since it is possible to group stimuli in almost any conceivable manner and to classify and subclassify them indefinitely, it is strictly true that the number of attitudes which any given person possesses is almost infinite. This result is statistically as well as psychologically absurd. Exactly the same absurdity and the same obstacle to research is offered by those definitions of attitude which conceive them merely as *verbal expressions* or as indications of assent to or dissent from particular verbal expressions. The number of possible verbal combinations is, of course, infinite and the number of attitudes must on this basis likewise be so. It is clear then that those who have defined attitudes in the above two ways have not meant exactly what

they have said. They have really intended to indicate not the actually discriminable tendencies to overt action or the verbal-response patterns but certain discernible *groups* of social responses. Within each group a family resemblance of the various responses is assumed, and each group of attitudes is supposed to show some distinguishable difference from every other group. If the analysis is pressed far enough, this turns out to mean that the attitude is a habit sufficiently compact and stable to be treated as a unit. It will, of course, be recognized that variations of an individual's response within this sphere of a given "attitude," together with differences between each attitude and the next, are involved. If it is my "attitude" to regard the eating of starches as a dietetic monstrosity, either an inherited dislike for such foods or a bundle of acquired tendencies directed towards bread, potatoes, rice, etc., must be assumed to exist. Whether we take the attitude therefore as an entity innate or learned, it is in either case not an inflexible and rigid element in personality (if, in fact, any such elements exist), but rather a certain *range within which responses move*.

On this basis one of our cardinal problems is to find whether social attitudes, in this sense, can be shown to be measurable, and if an affirmative answer is forthcoming, a serious attempt must be made to justify the separation of one attitude from others. For if there are no family differences between attitudes, there is simply *one* infinite series of attitudes. On this basis the measurement of attitudes could never mean anything more than the determination of an amount of some one tendency present, but what this one tendency actually is could never be defined.

Perhaps vaguely realizing this difficulty, many investigators in the field of attitude measurement have assumed a sharp distinction between different families of attitudes, and perfect resemblance or indeed identity between the members of any given family of responses. Take, for example, the frequent raising of questions as to whether there is a relation between internationalism and economic liberalism. Such terminology is based upon two assumptions: first, that there is an entity to be called internationalism, the boundaries of which can be so defined as to separate it from economic liberalism and permit one to tell which is which—otherwise the state-

ment is tautological—, second, it is assumed that *within* the sphere of internationalistic responses there is good prediction from one response to the next. Internationalism is treated as an entity such that the defining of a person's internationalism on a given issue would define his internationalism on all others.

We have not been concerned up to the present, to define the term "attitude," since the logic of the above argument seems to be the same when applied to either of two prevailing definitions of attitude. Contemporary definitions cluster about two chief conceptions: first, that attitudes are dispositions toward overt action; second, that they are verbal substitutes for overt action. The former usage seems to the present writer to be preferable.\* The verbal declarations of opinion and attitude are regarded as an *indirect method of measuring dispositions* which are most easily signified and expressed in verbal form. It is desirable, however, to point out that the same considerations regarding specificity and generality apply to the verbal declarations as to other social habits. The same problem of group factors and special factors appears even when dealing with paper and pencil behavior.

Consequently whether attitudes be defined in terms of underlying dispositions toward overt action or in terms of mere verbal declarations, investigators have proceeded almost without question upon the assumption that social attitudes are grouped in patterns or clusters. A series of verbal propositions dealing with the same general social issue are assumed to be more or less equivalent, or at least to be closely related so as to permit prediction from a knowledge of a subject's attitude on one issue to the same subject's attitudes on other aspects of the same issue. Similarly, overt behavior favorable to or in opposition to a racial, national, religious, or economic group, is assumed to have some predictive value in relation to conduct in the future when the same group or class is again involved. Speaking generally, attitudes may be conceived as "clustered" or linked together; a general pro-Japanese attitude, for example, may show itself in a series of pro-Japanese

\* If an exact definition be regarded as useful, the following (*Century Dictionary and Cyclopaedia*, 1904-1906, I, 374) is preferred: "relation of persons viewed as the expression of, or as affecting, feeling, opinion, intentions, etc."

verbal declarations or a series of pro-Japanese overt acts. In statistical language, a *group factor* is assumed at the outset. Now just as the student of mental organization recognizes group factors which vary greatly in their magnitude and significance, so the student of attitudes tends to think of relatively important group factors occurring in some attitude areas and relatively unimportant ones in others. Here, however, the agreement among the investigators ends. In the few short years during which any serious statistical sophistication has existed in attitude research, much acrimony has been evident between those who are concerned to regard the group factors as slight and unimportant, and those who regard them as large and significant. Thus we find Bain (4) insisting that attitudes cluster together only when they constitute alternative verbalizations of the same conative tendency in relation to a specifically defined social situation. The various studies in the specificity of character traits have led to a widespread and, it must be admitted, highly uncritical acceptance of the view that verbal propositions have no predictive value for anything else in life except similar verbal propositions confronted by the same subject under the same circumstances. On the other hand, an excessively general view of character traits is naively assumed with almost equal frequency by writers who have discussed radicalism and conservatism, some of whom have defined the characteristics of radicals and conservatives, suggesting evidence as to the psychological mechanisms making for such general trends. Thus as early as 1924, H. T. Moore (20) defined five basic psychological characteristics predisposing towards radicalism as contrasted with conservatism, and as recently as 1929 and 1930 G. W. Allport (2, 3) has argued for a cluster of traits, chiefly intellectual, which make for radicalism on political, economic, and other social issues.

It seems to be high time to cut through the statistical confusion which has resulted from the whole specificity-generality argument, to bring out the actual points of disagreement which separate these rival groups of psychologists, to make clear the statistical assumptions involved in all such methods of reasoning, and above all to test empirically in an extensive way the actual coherence or clustering of attitudes on a variety of public issues.

### III. PROCEDURE

The project conceived in 1929 by Gardner Murphy and the present writer aimed first of all to present a wide array of problems having to do with these five major "attitude areas": international relations, race relations, economic conflict, political conflict and religion. The attitude areas best covered in the questionnaire are those of race relations, international relations, and economic conflict. It was our conviction that very high specificity would exist among the elements used in our questionnaire, except insofar as the questions clearly dealt with the same issues. On the basis of results obtained by C. W. Hunter (15), for example, in the field of Negro-white relations, it was believed that attitudes toward segregation, toward eating with the Negro and toward lynching would be independent, and that in general any one specific attitude toward the Negro would bear no clear relation to the attitudes on other issues. This, of course, does not mean that we expected to obtain the rather ideal specificity which would be indicated by a zero reliability (by the split-half method) when comparing pro-Negro attitudes on some items with pro-Negro attitudes on others. It does mean, however, that we expected fifteen attitudes dealing with nearly fifteen different issues in which the Negro is involved, to give exceedingly low split-half reliability, and that in order to obtain anything like a true "scale" for the measurement of such a thing as pro- or anti-Negro feeling, one would have to employ dozens or even scores of questions. In the same way, the specificity of opinions on international and economic issues was assumed to be so great that an enormous number of items would have to be used in order to give a satisfactory reliability for any scale which could be constructed. It is important to emphasize these expectations at the outset, since the results, although in some respects confirming our belief, have in general yielded remarkably clear-cut group factors, group factors which in the light of the small number of items used must be regarded as having, for the student subjects used, a status as clear as that enjoyed by group factors of immediate memory, verbal

ability, and the like to which a number of statisticians have called attention.

The method by which the questionnaire was constructed was as follows. Having determined to study intensively the matter of international, inter-racial and economic attitudes, and, to a minor degree, political and religious attitudes, among large numbers of college students at typical American universities, a survey was made of the questionnaires already administered by other psychologists for these purposes. Among those which proved especially helpful were those of G. B. Neumann (23), C. W. Hunter (15) and R. W. George (9). In addition, about two hundred newspapers and magazines were rapidly surveyed during the autumn of 1929, declarations of opinion being culled for consideration, special emphasis being given to the more dogmatic types of opinion frequently found in editorials. A small number of questions were included from books, addresses and pamphlets, and a number were made up by the experimenters. Wherever it was possible to use questionnaire material which had previously been extensively tried out, and where, in a sense, "norms" were available, we preferred to use the questions exactly as they stood. In a few cases, it was necessary to abbreviate and simplify the questions in order to make sure that only one issue was involved and that ambiguity was avoided. In those instances in which we made up our own questions, we sought to emphasize simplicity, clarity, and brevity.

Without exception, the questions were presented in such a form as to permit a "judgment of value" rather than a "judgment of fact." Phrases such as "The United States should," or "We ought to," or "No man should be allowed" constantly reappeared. In a few instances it may seem on first inspection that a question has to do with a question of fact, but closer analysis will reveal the highly arbitrary character of such "facts." Perhaps the least desirable of all the questions used was the following: "Is war at present a biological necessity?" Such a question appears to many minds to be categorically a factual one; for example, from a neo-Malthusian point of view it may be regarded as capable only of an affirmative answer. The term "necessity," however, refers here more to the student's attitudes toward various wants than to any of those types of necessity which are discussed by physicists or

logicians. This is not offered in defense of the use of this particular item, which is regarded as one which should have been omitted; this explanation is offered only to make clear that at least in the great majority of cases and, we hope, in all, the inquiry has to do with the wants, desires, conative dispositions of the subjects, not with their opinions regarding matters of fact.

One further generalization may be offered regarding the plan underlying the choices of questions. Since value judgments are required, it was conceived that every issue might be presented in such a way as to allow the subject to take sides as between two clearly opposed alternatives. Furthermore each issue was so drawn that two conflicting groups of persons were either named or implied, and the subject allowed to affiliate himself with one or with the other group. In the struggle of the Negro, for example, to attain economic, political, or social equality, where the white man resists such equality, the subject has an opportunity either to take or to refuse to take the Negro standpoint. In the case of the conflict of relatively unfavored economic groups against those who enjoy special opportunities, and in the case of weaker nations which demand greater territorial or economic expansion than is at present permitted them, the same opportunity to ally oneself with one or the other of two opposing factions is involved. Again, it is not asserted that we have in all cases succeeded in framing an ideal "conflict issue." We would urge, however, that the great majority of conflict issues covered in our survey are empirically important issues, in which students at American universities actually do have opportunity to take sides and with regard to which the great majority have heard repeated discussions; secondly, that the results constitute in themselves an empirical check on the degree of success with which a tendency to take the side of a given group does enter into our questions in such a way as to be empirically measured. High specificity would have meant, among other things, that we had certainly failed in finding any general pro-Negro, anti-Japanese, etc. attitudes, and would have led to statistical difficulties of various sorts in handling incoherent masses of unrelated material. The clear-cut generality of certain attitudes, such as pro-Negro, internationalism, etc., shows that it is precisely in the field of *affiliation*

With or against certain social groups that the most definite results are obtained.

Through collaboration with instructors, the attitudes tests were given to undergraduates (chiefly male) in nine universities and colleges extending from Illinois to Connecticut and from Ohio and Pennsylvania to Virginia. (The names of the institutions cannot appropriately be printed here, only the Columbia College data, Group D, being identified.) The total number of individuals participating was somewhat above 2000 but the data here intensively analyzed were derived from only 650\* persons. The attitudes test, called a SURVEY OF OPINIONS, was first given in the late fall of 1929 (to all groups except Group C and Group F which were given the test in 1931) and, by arrangement with instructors, a retest given 30 days later. Some items from the first test and many new items were included in this second test. The first test required on the average about 40 minutes and the retest a slightly longer time.

The kind of questionnaire material to be reported here falls into four main classes. In the first, questions were to be answered by a Yes, a question mark, or a No, as for example, "Do you favor the early entrance of the United States into the League of Nations?" YES ? NO. Next came a series of multiple-choice questions in which one of five possible answers was to be selected, for example: "Using the term 'armaments' to mean equipment devised for war rather than for police purposes, our policy should be to favor: (a) absolute and immediate disarmament of all nations, (b) rapid and drastic reduction of the armaments of all nations, (c) slow but steady reduction of all armaments, (d) maintenance for a long time of approximately the present military and naval strength of all the powers, (e) our free military and naval expansion unembarrassed by agreements with other nations." Third, there was a series of propositions to be responded to by the words (a) *strongly approve*, (b) *approve*, (c) *undecided*, (d) *disapprove*, (e) *strongly disapprove*, for example: "All men who have the opportunity should enlist in the Citizens Military

---

\* These 650 cases represent a random sample from seven of the groups comprised in the study. The data on two of the nine colleges have not as yet been analyzed. Of course, only those individuals were used for whom we have complete data.

Training Camps." Fourth, a series of abbreviated newspaper narratives about social conflicts, terminating in a sentence describing the *outcome* of this conflict, the student being asked to indicate his response to this outcome, for example: "A group of Japanese truck-farmers in Southern California, through their industry and lower standards of living, are able to undersell their American competitors. The American farmers insist that IT IS THE DUTY OF ALL WHITE PEOPLE TO PURCHASE ONLY FROM WHITE FARMERS." This last form of question makes use of the same set of five responses mentioned above, *strongly approve*, *approve*, *undecided*, *disapprove*, and *strongly disapprove*.

### THE ATTITUDE SCALES

The different scales presented here have been given their respective names merely for convenience in referring to them. The names given them seem the most plausible, but to avoid any "jingle fallacy" it should be recognized that the scales measure merely what the different statements included in them involve.

In the SURVEY OF OPINIONS the statements did not appear consecutively as shown in the following scales but were scattered among many statements having to do with other attitudes.

#### INTERNATIONALISM SCALE

The numbers in parentheses below or at the side of the different alternatives show the numerical value assigned to them in the final scoring. These numbers were, of course, not present upon the SURVEY OF OPINIONS when taken by the subject.

1. Do you favor the early entrance of the United States into the League of Nations?
 

YES	?	NO
(4)	(3)	(2)
2. Ought the United States to consult other nations in making her immigration laws?
 

YES	?	NO
(4)	(3)	(2)
3. Should the United States give naval demonstrations in the Pacific?
 

YES	?	NO
(2)	(3)	(4)
4. Is war at present a biological necessity?
 

YES	?	NO
(2)	(3)	(4)

5. Should the United States recognize the Soviet government?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (4) | (3) | (2) |
6. Should the Treaty of Versailles be reconsidered, with greater leniency given to Germany?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (4) | (3) | (2) |
7. Should the United States cancel a large part of the Allied war debt in return for concessions as to disarmament and economic reconstruction abroad?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (4) | (3) | (2) |
8. Should there be a national referendum on every war?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (4) | (3) | (2) |
9. Do you look with suspicion upon the idea of a Super-State as the future hope of international government?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (2) | (3) | (4) |
10. Is it an idle dream to expect to abolish war?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (2) | (3) | (4) |
11. Are you in sympathy with the movement for the outlawing of war?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (4) | (3) | (2) |
12. Should the United States enter the World Court?
- |     |     |     |
|-----|-----|-----|
| YES | ?   | NO  |
| (4) | (3) | (2) |
13. How much military training should we have?
- (a) We need universal compulsory military training. (1)
  - (b) We need Citizens Military Training Camps and Reserve Officers Training Corps, but not universal military training. (2)
  - (c) We need some facilities for training reserve officers but not as much as at present. (3)
  - (d) We need only such military training as is required to maintain our regular army. (4)
  - (e) All military training should be abolished. (5)
14. How large should our navy be?
- (a) We should maintain the "two-power standard" formerly maintained by Great Britain (i.e. ours should be as strong as any two others). (1)
  - (b) We should maintain a considerable margin over our nearest competitor. (2)
  - (c) We should share first place with another power as at present. (3)
  - (d) We should attempt only to maintain second place. (4)
  - (e) We should not be especially concerned with our rank as a naval power. (5)

MEASUREMENT OF ATTITUDES

15. Using the term "armaments" to mean equipment devised for war rather than for police purposes, our policy should be to favor:
- (a) absolute and immediate disarmament of all nations. (5)
  - (b) rapid and drastic reduction of the armaments of all nations. (4)
  - (c) slow but steady reduction of all armaments. (3)
  - (d) maintenance for a long time of approximately the present military and naval strength of all the powers. (2)
  - (e) our free military and naval expansion unembarrassed by agreements with other nations. (1)

16. All men who have the opportunity should enlist in the Citizens Military Training Camps.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(1)	(2)	(3)	(4)	(5)

17. The United States, whether a member or not, should co-operate fully in the humanitarian and economic programs of the League of Nations.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(5)	(4)	(3)	(2)	(1)

18. In the interest of permanent peace, we should be willing to arbitrate absolutely all differences with other nations which we cannot readily settle by diplomacy.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(5)	(4)	(3)	(2)	(1)

19. A person who loves his fellow men should refuse to engage in any war, no matter how serious the consequences to his country.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(5)	(4)	(3)	(2)	(1)

20. The United States should have the largest military and naval air fleets in the world.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(1)	(2)	(3)	(4)	(5)

21. We should be willing to fight for our country whether it is in the right or in the wrong.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(1)	(2)	(3)	(4)	(5)

22. We must strive for loyalty to our country before we can afford to consider world brotherhood.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(1)	(2)	(3)	(4)	(5)

23. Our country should never declare war again under any circumstances.

Strongly Approve	Approve	Undecided	Disapprove	Strongly Disapprove
(5)	(4)	(3)	(2)	(1)

ter leni-  
war debt  
ic recon-

ate as the

g of war?

(1)  
reserve  
train-  
(2)  
ers but  
(3)  
ired to  
(4)  
(5)

ormerly  
l be as  
(1)  
ver our  
(2)  
r as at  
(3)  
e.  
(4)  
ank as a  
(5)

24. Moving pictures showing military drill and naval manoeuvres should be exhibited to encourage patriotism.

Strongly Approve (1)	Approve (2)	Undecided (3)	Disapprove (4)	Strongly Disapprove (5)
----------------------------	----------------	------------------	-------------------	-------------------------------

### NEGRO SCALE

(Statement Number 4 was not included in the odd statements or the even statements when finding the split-half reliability.)

1. Would most negroes, if not held in their place, become officious, overbearing, and disagreeable?

YES (2)	? (3)	NO (4)
------------	----------	-----------

2. If you went into a cafeteria in a northern city, sat down, and then realized you were at the table with a negro, would you leave the table?

YES (2)	? (3)	NO (4)
------------	----------	-----------

3. Would you shake hands with a negro?

YES (4)	? (3)	NO (2)
------------	----------	-----------

4. Do you disapprove of the use of the term "nigger"?

YES (4)	? (3)	NO (2)
------------	----------	-----------

5. If you heard of a negro who had bought a home or a farm would you be glad?

YES (4)	? (3)	NO (2)
------------	----------	-----------

6. In a community in which the negroes outnumber the whites, under what circumstances is the lynching of a negro justifiable?

- |   |     |
|---|-----|
| (a) Never.  | (5) |
| (b) In very exceptional cases where a specially brutal crime against a white person calls for swift punishment. | (4) |
| (c) As punishment for any brutal crime against a white person.  | (3) |
| (d) As punishment for any gross offense (felony or extreme insolence) committed against a white person.         | (2) |
| (e) As punishment for any act of insolence against a white person.  | (1) |

7. How far in our educational system (aside from trade education) should the most intelligent negroes be allowed to go?

- |                                       |     |
|---------------------------------------|-----|
| (a) Grade school.                     | (1) |
| (b) Junior high school.               | (2) |
| (c) High school.                      | (3) |
| (d) College.                          | (4) |
| (e) Graduate and professional school. | (5) |

8. In a community where the negroes outnumber the whites, a negro who is insolent to a white man should be:

- |   |     |
|---|-----|
| (a) excused or ignored.   | (5) |
| (b) reprimanded.  | (4) |
| (c) fined and jailed.   | (3) |
| (d) not only fined and jailed, but also given corporal punishment (whipping, etc.). | (2) |
| (e) lynched.  | (1) |

vres should  
Strongly Disapprove (5)  
ents or the  
ficious, over-  
wn, and then you leave the  
arm would you  
whites, under able?  
ital crime ent. (5)  
a white (4)  
r extreme (3)  
st a white (2)  
ade education) o? (1)  
whites, a negro (5)  
oral punish- (4)  
(3)  
(2)  
(1)

9. All negroes belong in one class and should be treated in about the same way.
 

Strongly Approve (1)	Approve (2)	Undecided (3)	Disapprove (4)	Strongly Disapprove (5)
-------------------------	----------------	------------------	-------------------	----------------------------
10. Negro homes should be segregated from those of white people.
 

Strongly Approve (1)	Approve (2)	Undecided (3)	Disapprove (4)	Strongly Disapprove (5)
-------------------------	----------------	------------------	-------------------	----------------------------
11. 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 (5)	Approve (4)	Undecided (3)	Disapprove (2)	Strongly Disapprove (1)
-------------------------	----------------	------------------	-------------------	----------------------------
12. If the same preparation is required, the negro teacher should receive the same salary as the white.
 

Strongly Approve (5)	Approve (4)	Undecided (3)	Disapprove (2)	Strongly Disapprove (1)
-------------------------	----------------	------------------	-------------------	----------------------------
13. Practically all American hotels should refuse to admit negroes.
 

Strongly Approve (1)	Approve (2)	Undecided (3)	Disapprove (4)	Strongly Disapprove (5)
-------------------------	----------------	------------------	-------------------	----------------------------
14. No negro should be deprived of the franchise except for reasons which would also disfranchise a white man.
 

Strongly Approve (5)	Approve (4)	Undecided (3)	Disapprove (2)	Strongly Disapprove (1)
-------------------------	----------------	------------------	-------------------	----------------------------
15. In a community of 1,000 whites and 50 negroes, a drunken negro shoots and kills an officer who is trying to arrest him. THE WHITE POPULATION IMMEDIATELY DRIVE ALL THE NEGROES OUT OF TOWN.
 

Strongly Approve (1)	Approve (2)	Undecided (3)	Disapprove (4)	Strongly Disapprove (5)
-------------------------	----------------	------------------	-------------------	----------------------------

IMPERIALISM SCALE

1. Should the United States grant complete independence to the Philippines at an early date?
 

YES (4)	? (3)	NO (2)
------------	----------	-----------
2. Should our forces be withdrawn from Nicaragua?
 

YES (4)	? (3)	NO (2)
------------	----------	-----------
3. The policy of the United States in maintaining the possession of the Philippine Islands:
 

(a) is exactly what it should be.	(1)
(b) has on the whole more to be said for it than against it.	(2)
(c) has about as many good points as bad ones.	(3)
(d) is difficult to justify.	(4)
(e) is a disgrace to our country.	(5)

4. Economic exploitation of territories and colonies by great powers:
- |  |     |
|--|-----|
| (a) is totally unjustifiable.                                    | (5) |
| (b) has some justification, but is on the whole wrong.           | (4) |
| (c) has about as many unjustifiable aspects as justifiable ones. | (3) |
| (d) has some questionable aspects, but on the whole is right.    | (2) |
| (e) is entirely reasonable and right.                            | (1) |
5. An American doing business in China should be willing to abide by Chinese law.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (5)              | (4)     | (3)       | (2)        | (1)                 |
6. We should use military force in South America whenever needed to protect American investments.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (1)              | (2)     | (3)       | (4)        | (5)                 |
7. In case of severe political and economic upset in China, western nations should maintain sufficient military forces in China to protect all interests previously acquired by their citizens.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (1)              | (2)     | (3)       | (4)        | (5)                 |
8. We should be willing to let American investments in China be lost rather than be drawn into armed conflict with China.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (5)              | (4)     | (3)       | (2)        | (1)                 |
9. Our army and navy should be used when necessary to defend American property rights in Mexico.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (1)              | (2)     | (3)       | (4)        | (5)                 |
10. An American committing a crime in Shanghai should be tried by a Chinese judge.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (5)              | (4)     | (3)       | (2)        | (1)                 |
11. We should co-operate as fully as possible with Latin-American countries, treat them as equals, and stop regarding ourselves as their leaders and protectors.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (5)              | (4)     | (3)       | (2)        | (1)                 |
12. As a result of inflammatory press dispatches, mobs in a small Latin-American country have repeatedly attacked United States flags and torn them to shreds. The United States citizens feel that their lives are in danger. MARINES ARE SENT TO PROTECT THE LIVES AND PROPERTY OF THESE CITIZENS.
- |                  |         |           |            |                     |
|------------------|---------|-----------|------------|---------------------|
| Strongly Approve | Approve | Undecided | Disapprove | Strongly Disapprove |
| (1)              | (2)     | (3)       | (4)        | (5)                 |

## IV. RESULTS

### 1. *The Sigma Method of Scoring*

In order to compare one type of statement with another such as the "multiple choice" with the "strongly approve," it was necessary to devise some technique whereby they might be made comparable. In attempting to work out such a technique, it was noticed that a great number of the five-point statements, i.e. the "multiple choice" or "strongly approve" statements (in each case the subject being offered five alternatives from which to choose), yielded a distribution resembling a normal distribution.

TABLE I  
PERCENTAGE OF INDIVIDUALS CHECKING THE DIFFERENT ALTERNATIVES.  
COMPUTED FROM A SAMPLE OF 100 CASES, ALL MALE, FROM A SINGLE  
UNIVERSITY

		MULTIPLE CHOICE STATEMENTS				
		<i>Alternatives</i>				
<i>Scale</i>	<i>Statement Number</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
Negro	7	1	1	3	8	87
Negro	8	29	42	26	3	0
Imperialism	3	11	43	27	15	4
		STRONGLY APPROVE STATEMENTS				
		<i>Alternatives</i>				
		<i>Strongly Approve</i>	<i>Un- Approve</i>	<i>Dis- decided</i>	<i>approve</i>	<i>Strongly Disapprove</i>
International-	16	13	43	21	13	10
Negro	9	3	17	14	44	22
Imperialism	5	32	52	10	5	1
Negro	10	24	49	17	7	3
Imperialism	6	10	27	17	35	11

Table I shows some typical distributions obtained with the five-point statements. These percentages are based on a sample of 100 cases, all male, from one university. Of the two distributions which were quite skewed, number 7 of the Negro scale, was the more skewed. The other was number 6 of the Negro scale. It is interesting to note that if a group of Southern students are included in these distributions they become

less skewed. Thus, for statement number 7 of the Negro scale the percentages for 100 male students from a college in Virginia are, respectively, 4, 3, 17, 18 and 58. Statements number 5 and 6 of the Imperialism scale are illustrative of a slight bi-modality which was found in a few of the "strongly approve" type of statements.

On the basis of this experimental evidence and upon the results of others (8, pp. 542-548, 28, pp. 71-91), it seems justifiable for experimental purposes to assume that attitudes are distributed fairly normally and to use this assumption as the basis for combining the different statements. The possible dangers inherent in this assumption are fully realized. This assumption is made simply as part of an experimental approach to attitude measurement. It is a step which it is hoped subsequent work in this field will either make unnecessary or prove justifiable. Perhaps this assumption is not correct; its correctness or incorrectness can best be determined by further experiment.

The percentage of individuals that checked a given position on a particular statement was converted into sigma values. This was done for each of the five-point statements which in our opinion had to do with internationalism. Table 22 of Thorndike's tables (30) greatly facilitated this calculation. These tables assume that one hundred per cent of the cases fall between  $-3$  and  $+3$  sigma. The values given in the table are the average sigma values of intervals represented by the stated percentages, the origin considered to be at the mean. The sigma deviations were always taken from the mean and the positive value was assigned to the end which seemed to favor internationalism, the negative being assigned to the end which favored nationalism. To avoid using negative values the arbitrary zero may be placed at  $-3$  sigma rather than at the mean. These signs were designated in an arbitrary fashion and then verified objectively (see pages 48-52). The sigma values were computed from percentages obtained from a sample of 100 cases, all male, selected from one particular university. Table II shows the percentage of individuals checking each of the different alternatives and the corresponding sigma values for statement number sixteen of the Internationalism scale.

The statements selected were checked for internal consist-

TABLE II  
DATA FOR STATEMENT NUMBER 16 OF THE INTERNATIONALISM SCALE

<i>Alternative</i>	<i>Strongly Approve</i>	<i>Approve</i>	<i>Un-decided</i>	<i>Dis-approve</i>	<i>Strongly Disapprove</i>
Percentage checking	13%	43	21	13	10
Corresponding sigma value	-1.63	-.43	+.43	+.99	+1.76
Corresponding 1 to 5 value	1	2	3	4	5

ency or "clustering," by finding the reliability, using odd statements vs. even statements. The fourteen five-point statements used yielded moderately high reliabilities when tried on three different groups with between 30 and 35 subjects in each group. Two of these groups were from the same university, the third was from another university in an entirely different geographical area. These results indicate a "cluster" or attitude variable which we are justified in treating as a unit, so far as these three groups are concerned. The reliabilities obtained for these groups are given in Table III. These results and the following considerations seem to justify the statement that the sigma scoring technique is the most satisfactory now available for attitude measurement. It not only seems to avoid many of the shortcomings of existing methods of attitude measurement, but at the same time retains most of the advantages present in methods now used.

TABLE III  
RELIABILITY COEFFICIENTS—SIGMA SCORING METHOD—FOURTEEN STATEMENTS DEALING WITH INTERNATIONALISM  
Odds vs. Evens (7 items vs. 7 items)

<i>Group</i>	<i>N</i>	<i>Raw</i>	<i>Corrected</i>
A	30	.76	.86
B	32	.79	.88
F	33	.75	.86

In the first place, the sigma scoring method meets the requirement stated by Thurstone (38, p. 56):

"Ideally, the scale should perhaps be constructed by means of the voting only. It may be possible to formulate the prob-

lem so that the scale values of the statements may be extracted from the records of actual voting. If that should be possible, then the present procedure of establishing the scale-values by sorting will be superseded."

Further, it avoids the difficulties encountered when using a judging group to construct the scale. A number of these difficulties have been pointed out by Rice (27). The following quotation deals with one of the major shortcomings of any technique employing a judging group (27, pp. 190-191):

"The difficulties of building scales similar to Thurstone's and of applying them to the measurement of the attitudes of social groups, become increasingly difficult once we leave the classroom, the discussion club and the other small, comparatively infrequent and highly selected groups that enjoy having experiments tried upon them. Such groups already have developed ways of making their attitudes articulate. It is the more numerous work-a-day groupings of society, which are inaccessible to his controlled measurements, about whose attitudes the social scientist is in the most need of information. Students may be required, good natured academicians may be cajoled, and sundry needy persons may be paid to sort cards containing propositions into eleven piles. But it is difficult to imagine securing comparable judgments, or satisfactory measurements in the final application, from bricklayers, business men, Italian-Americans, nuns, stevedores, or seamstresses. And, unless the scale itself is based upon equal-seeming differences to a random sample of the group which is to be measured, its validity—the degree to which it measures that which it purports to measure—becomes open to question."

Another decided advantage of the sigma technique is that it yields reliabilities as high as those obtained by other techniques, with fewer items. This is possible because it uses an approach to the problem somewhat different from that conventionally used. Previously attempts have been made to find the scale value of each particular statement along a continuum; a person's score being then determined by the scale value of the statements that he accepts. In this study, however, each statement becomes a scale in itself and a person's reaction to each statement is given a score. These scores are then combined by using a median or a mean. Eggan's study reported by Thurstone (35) lends further evidence to support the method presented here.

In contemplating this method of measuring attitudes it is well to realize that the stronger the generic set toward one extreme or the other extreme of an attitude continuum, the more it influences the specific reactions. When the generic set is not strong then the specific items themselves largely determine the reaction. In the latter case, however, the reaction is seldom very intense but rather mildly pro or con. That is, the individual's reactions, so far as that particular attitude is concerned, do not deviate widely from the average.

The sigma technique also yields scores the units of which are equal throughout the entire range. Likewise, the same kinds of measures can be obtained with it as are obtained with other techniques now in existence (38). Thus it is possible to obtain the most typical measure of an individual's attitude and also the range or dispersion of his attitude.

Needless to say the construction of an attitude scale by the sigma method is much easier than by using a judging group to place the statements in piles from which the scale values must be calculated.

Among the excellent characteristics of Thurstone's method of attitude construction (38) are the objective checks which he has devised for ambiguity and irrelevance. Similar objective checks can be applied to the sigma technique, if desired. The application of these objective checks are fully discussed on pages 48-52.

It is interesting to note that the scores on the international statements using the sigma technique correlate  $+0.67$  with the Thurstone-Droba War scale (6) (data from Group F). When corrected for attenuation, this becomes  $+0.77$ . This relationship is present even though the statements on the whole are quite dissimilar.

## 2. *The Simpler Method of Scoring*

Although the sigma technique seemed to be quite satisfactory for the intended use, it was decided to try a simpler technique to see if it gave results comparable with the sigma technique. If it did, the simpler method would save considerable work in a general survey type of study of this kind. The simpler technique involved the assigning of values of from 1 to 5 to each of the five different positions on the five-point statements. The ONE end was always assigned to the nega-

y be extracted  
ld be possible,  
scale-values by

ed when using  
A number of  
(27). The fol-  
hortcomings of  
(pp. 190-191):

to Thurstone's  
the attitudes of  
ce we leave the  
small, compara-  
that enjoy hav-  
ps already have  
ulate. It is the  
dety, which are  
about whose at-  
of information.  
emicians may be  
aid to sort cards  
t it is difficult to  
atisfactory meas-  
layers, business  
or seamstresses.  
ual-seeming dif-  
ich is to be meas-  
sures that which  
stion."

technique is that  
ed by other tech-  
because it uses an  
it from that con-  
been made to find  
ent along a con-  
ained by the scale  
n this study, how-  
elf and a person's

These scores are  
n. Eggan's study  
vidence to support

tive end of the sigma scale, and the FIVE end to the positive end of the sigma scale. (See Table II.)

After assigning in this manner the numerical values to the possible responses, the score for each individual was determined by finding the average of the numerical values of the positions that he checked. Actually, since the number of statements was the same for all individuals, the sum of the numerical scores rather than the mean was used. The reliability of odds vs. evens for this method yielded essentially the same values as those obtained with the sigma method of scoring. The scores obtained by this method and the sigma method correlated almost perfectly as will be seen in Table IV.

TABLE IV  
COEFFICIENTS OF CORRELATION BETWEEN SCORES OBTAINED BY THE SIGMA METHOD, THE 1-5 METHOD, AND THE 1-7 METHOD

Group	N	INTERNATIONAL (15 Statements)			NEGRO (10 Statements)
		$\sigma$ vs. 1-5	1-5 vs. 1-7	$\sigma$ vs. 1-7	$\sigma$ vs. 1-5
A	30	.991	.990		.987
B	32	.995	.993		.992
F	33	.995	.990	.997	

The same results were obtained when the values of 1, 3, 4, 5 and 7 were assigned to the different positions corresponding respectively to 1, 2, 3, 4 and 5. In the former case, it will be noted that the extremes were given slightly greater weight. This method likewise correlated very highly with the sigma method and with the 1 to 5 method as shown in Table IV.

These results seem to justify the use of the simpler methods of scoring since they yield almost identical results with the sigma method and similarly do not involve any of the errors likely to be present in any technique in which experts, judges, or raters are used.

Furthermore, the 1 to 5 method has the additional advantage over the sigma method for this particular study in that it permits us to combine the reactions on the three-point statements in the SURVEY OF OPINIONS with those on the five-point statements and thus get a more reliable measure of the student's attitude. A three-point statement, it will be recalled, is one to be answered by a YES ? NO.

Three methods of scoring the three-point statements were tried. The first assigned values of 2, 3 and 4 to the alternative responses, the 4 always being assigned to the particular response which seemed to us most "internationalistic." The question mark, of course, always had a value of 3. The second method involved the assigning of values of 1, 3 and 5 to the particular responses. The third method involved the assigning of positive and negative signs in the same manner as was done with the sigma technique in the case of the five-point statements, but here the numerical value assigned was the percentage of individuals who checked the opposite position. In this manner the responses were weighted in terms of typicality and atypicality much as was done in the sigma method. For example, if 60% of the individuals checked YES, and 25% checked NO (15% checking ?), and if the YES is the more international—hence designated positive—the score value of the YES is plus 25 and of NO minus 60. In this case the question marks were ignored.

It was found that all three methods yielded essentially the same results. Since this was the case, it was decided in this study to use the simpler methods throughout.\*

Using the 2-3-4 method for the three-point statements and the 1 to 5 method for the five-point statements, a scale was constructed for measuring attitudes toward internationalism. In constructing this scale, not only was the reliability determined in essentially the same manner as was done in the sigma technique, but the criterion of internal consistency also was applied. (See pages 50-52 for a discussion of the criterion of internal consistency.) It was discovered that, while there was considerable internal consistency between all the statements used, there appeared to be two groups of statements which showed greater internal consistency than inter-group consistency. These two groups of statements when examined indicated that the one group consisted of statements having to do more specifically with imperialism, while the other group had to do more with other problems in the field of international relations. Consequently, a separate scale was con-

\* It will be noted that the writer began this inquiry with a suspicious attitude toward the simple computations used in rating scales, and adopted these simple procedures only in the light of evidence showing that the simpler methods gave the same results as the elaborate.

ie positive  
ues to the  
was deter-  
ues of the  
er of state-  
of the nu-  
reliability  
the same  
of scoring.  
method cor-

THE SIGMA

NEGRO  
Statements)

$\sigma$  vs. 1-5

.987

.992

of 1, 3, 4, 5  
responding  
e, it will be  
ter weight.  
the sigma  
ble IV.  
er methods  
s with the  
the errors  
rts, judges,  
onal advan-  
y in that it  
point state-  
on the five-  
sure of the  
be recalled,

structed for each of these attitudes. In constructing the Imperialism scale the high or FIVE value was always assigned to the anti-imperialistic point of view; hence, the higher the individual score, the *less* imperialistic the individual.

In a similar manner, using all the statements in the SURVEY OF OPINIONS which had to do with Negroes, a scale was constructed for measuring attitudes toward the Negro. Table X on page 50 shows the internal consistency obtained with the Negro scale. The higher scores are indicative of an attitude favorable to the Negro. Some may question whether the numerical values assigned to the different responses of a particular statement are valid, asking if the numerical values ought not be the reverse in order to conform to "favorable." But it is noteworthy that in every case the assigning of the numerical values to the different alternative responses was found to be consistent when checked by item analysis or the criterion of internal consistency. Among the 87 statements used (24 in the Internationalism scale, 15 in the Negro scale, 12 in the Imperialism scale, and 36 in the Thurstone-Droba scales—see pages 33-35) not once was it necessary to reverse the numerical values arbitrarily assigned by the experimenter to the alternative responses. If the numerical values assigned to the alternative responses of a statement, for example, of the Negro scale are questioned, then the values assigned to all of the statements must be questioned. Judging by the internal consistency obtained, the experimenter is either right on all the statements or wrong on all of them. To call "favorable" to the Negro the opposite alternative of that selected seems to be contrary to ordinary usage.

The degree of interrelationship between our three scales may be briefly indicated for a group of 100 Columbia College students (Group D). The raw correlation coefficient between the Negro and Internationalism scales is  $+.40 \pm .06$ , corrected  $+.46$ ; that between Negro and Imperialism, raw  $+.34 \pm .06$ , corrected,  $+.40$ ; that between Imperialism and Internationalism, raw,  $+.63 \pm .04$ , corrected,  $+.70$ . Material from three other universities has yielded very similar results.

### 3. *The Reliability of the Scales*

The reliability coefficients for the three scales with the different groups are given in Table V. A table of the probable

TABLE V  
RELIABILITY COEFFICIENTS (ODDS VS. EVENS)

Group	N	INTERNATIONALISM SCALE 24 Items (12 vs. 12)		IMPERIALISM SCALE 12 Items (6 vs. 6)		NEGRO SCALE 15 Items (7 vs. 7)	
		Raw	Corrected	Raw	Corrected	Raw	Corrected
A	30 (9 female)	.73	.84	.77	.87	.79	.88
B	32 (10 female)	.79	.88	.73	.85	.83	.91
C	55 (male only)	.79	.88	.67	.80	.72	.84
D	100 (male only)	.81	.90	.79	.88	.73	.84
E	85 (male only)	.68	.81	.73	.84	.70	.82
F	33 (male only)	.78	.88	.85	.92	.72	.84
G	100 (male only)	.82	.90	.71	.83	.74	.85
H	92 (male only)	.74	.85	.73	.84	.65	.79

acting the Im-ways assigned the higher the idual. s in the SUR-egroes, a scale rd the Negro. tency obtained cative of an at- on whether the onses of a par- umerical values to "favorable." ssigning of the responses was analysis or the e 87 statements the Negro scale, Thurstone-Droba ssary to reverse he experimenter l values assigned r example, of the assigned to all of g by the internal ther right on all all "favorable" to selected seems to

our three scales Columbia College coefficient between  $40 \pm .06$ , corrected  $1$ , raw  $+ .34 \pm .06$ ,  $1$  and International-aterial from three results.

ee scales with the ble of the probable

errors of correlation coefficients is given in Table VI. Groups A, B and H are from the same university. The other groups are all from different universities. With the exception of Groups A and B only male subjects were used. Only men were used in order to make the groups more homogeneous.

In general, the Internationalism scale seems to have a somewhat higher reliability than the other two scales. This is not surprising in view of its greater length. It is not, however, as reliable as might be expected when its length is compared

TABLE VI  
PROBABLE ERRORS OF THE COEFFICIENT OF CORRELATION FOR VARIOUS  
NUMBERS OF MEASURES (N) AND FOR VARIOUS VALUES OF  $r$

<i>Number of Measures</i>	<i>Correlation Coefficient (r)</i>						
	(N)	0.65	0.70	0.75	0.80	0.85	0.90
30		071	063	054	044	034	023
50		055	049	042	034	027	018
70		047	041	035	029	022	015
100		039	035	029	024	019	013

with the length of the other scales. The most plausible reason for this is that it has a much greater proportion of three-point statements than the other scales. The Internationalism scale has twelve three-point statements and twelve five-point statements, the Negro scale has five of the former and ten of the latter, and the Imperialism scale has respectively, two and ten. On the whole the three-point statements were found to be less satisfactory than the five-point statements, but this might well be the result of a chance selection of statements in the construction of the SURVEY OF OPINIONS rather than a function of the number of alternatives presented.

In general it would seem desirable to set .90 as a minimum reliability coefficient for a test, due consideration, of course, being given to variability. It is regrettable that there were not sufficient statements in the SURVEY OF OPINIONS to make each scale long enough to yield this reliability.

In the Negro scale statement number 4 was omitted when calculating the split-half reliability. This was done to have

an even number of statements in the two halves of the scale so as to permit the use of the usual form of the Spearman-Brown formula when calculating the reliability coefficients for the entire test.

It is surprising that the 30-day retest reliabilities are so high when it is recognized that this coefficient is the result first, of the reliability of the test itself and, second, of the changes in the attitudes of the individual subjects during the interval. Apparently one is justified in concluding that during this 30-day interval the attitudes of the different subjects were relatively constant in view of the fairly high reliability coefficients that were obtained. In a few cases there were marked shifts in a given direction on the part of a particular individual. This generally was due to some event which had occurred between the first and second tests which caused him to change his attitude considerably. For example, one student in Group F was in a course in history in which considerable information concerning Latin-America was presented. Between the first and second tests information was given him which he interpreted as justifying imperialism on the part of the United States in Latin-America. This resulted in his shifting a total of nine points toward a more imperialistic point of view. Other similar instances could be presented, such as that of the individual who was severely beaten by a Negro between the first and second tests. If cases of this kind are omitted, the 30-day retest reliability is of course much higher.

In part, at least, the high reliability of the 30-day retest may be due to the fact that there is some retention of answers on the part of the individuals concerned. When asked, however, they invariably reported an inability to recall their specific answers on the previous test. The fact that they were unable to recall does not mean that the subject would not show some retention by the "saving method." The inability to recall one's answer after a period of thirty days suggests that a 30-day retest often may be helpful to determine the honesty of the individual answering the first test. This will not determine every case of insincerity in answering the test because in some cases a person might answer, on both the original test and on the retest, not in terms of his true attitude but in terms of what he thinks his attitude is expected to be.

VARIOUS  
OF ?

.85	0.90
34	023
27	018
22	015
19	013

ible reason  
three-point  
ialism scale  
point state-  
l ten of the  
wo and ten.  
id to be less  
s might well  
in the con-  
than a func-

s a minimum  
n, of course,  
t there were  
PINIONS to  
ity.  
mitted when  
done to have

Groups  
groups  
tion of  
ly men  
ous.  
a some-  
is not  
owever,  
ompared

4. *The Validity of the Scales*

In any discussion of the validity of attitude scales of the kind presented here it might be well to emphasize that at present we are dealing only with verbal behavior and claim nothing more than the importance of the verbal reactions. Ultimately it is to be hoped that the relationship between the verbal behavior expressed on an attitude scale and other more overt forms of behavior may be examined and determined, but at present we are concerned with verbal behavior only. After all, the verbal reactions studied here are related to problems in reference to which the majority of our reactions in everyday life are verbal. We declare ourselves in favor of one issue and opposed to another, and such declarations are socially accepted as symbols for overt acts. In many cases it would seem reasonable to conclude that since our daily behavior in these areas is largely verbal, the verbal responses would be valid indices of other habits.

Attention, moreover, should be called to the work of Thurstone (38), Neumann (23), Kulp and Davidson (17), Porter (25), and Stouffer (29) in showing the validity of attitude scales in measuring attitudes on social questions. It is perhaps striking that the individual who (in 1929) was the only person to make the highest possible score on the Internationalism scale, who was one of two that made the highest possible score on the Negro scale and who made the highest (anti-) Imperialism score yet recorded, is a student who has recently taken a very prominent part in radical activities on the Columbia Campus.

TABLE VII  
MEANS\* AND STANDARD DEVIATIONS OF THE DIFFERENT GROUPS ON THE  
THREE SCALES

Group	N	Internationalism		Imperialism		Negro	
		M	S.D.	M	S.D.	M	S.D.
A	30	75.52	8.98	35.63	7.34	56.13	6.31
B	32	73.09	11.80	34.65	7.10	55.53	6.77
C	55	82.40	12.40	40.25	5.95	55.87	5.12
D	100	80.24	12.05	37.99	6.98	59.40	6.09
E	85	67.27	9.35	33.41	6.56	54.22	6.23
F	33	79.30	9.43	39.03	6.98	60.21	5.27
G	100	74.76	12.04	35.15	7.11	56.44	6.23
H	92	73.51	10.31	36.67	7.04	54.73	5.86
I	123	—	—	—	—	46.35	7.36

Table VII shows the means (M) and the standard deviations (S.D.) of the different groups on the different scales. It is interesting to notice, relative to validity, that on the Negro scale the lowest of the northern college groups, Group E, has a mean 7.87 points above that of the college in Virginia. This difference (7.87) when divided by the sigma of the difference (.95) is 8.3 which indicates that it is a reliable difference.

Of course, it cannot be sufficiently emphasized that the reactions on an attitude test are no more meaningful than the situation in which the attitude test was given. If the situation is such as to elicit the honest cooperation of the subject, so that he will be likely to state his own attitude and not the attitude that he thinks is expected of him or some other equally fictitious attitude, we can feel that we have a valid measure of his attitude. The danger of not having the full cooperation of the subject cannot be overemphasized in the present promiscuous use of attitude tests. It is significant that the reliabilities of the different scales tend to be higher, when allowance is made for variability, in those groups where there was more reason to expect the cooperation of the student.

5. *Comparison of the Simpler Method with the Thurstone Method of Scoring*

Two groups, C and F, were given the Thurstone-Droba War scale (6) as well as the SURVEY OF OPINIONS. Table VIII shows the reliability coefficients obtained for the Thurstone scale and for the Internationalism scale, derived from the SURVEY OF OPINIONS, for these two groups. The reliability coefficients of the Thurstone test, obtained by correlating Form A against Form B, was respectively .78 and .74 for the two groups. The reliability for the two forms combined, as determined by the Spearman-Brown formula, becomes .88 and .85, respectively. The same reliability is obtained by the present Internationalism scale with 24 items as is obtained by combining both forms of the Thurstone-Droba scale with a total of 44 items. Thus using the method here described, a measure of a person's attitude as reliable as that obtained by the Thurstone method is secured by asking him to react to one-half as many items. The coefficients of correlation between the Internationalism scale and the Thurstone-Droba scale are also given in Table VIII.

ude scales of the  
asize that at pres-  
r and claim noth-  
al reactions. Ul-  
iship between the  
le and other more  
l and determined,  
oal behavior only.  
re related to prob-  
f our reactions in  
selves in favor of  
h declarations are  
s. In many cases  
at since our daily  
ie verbal responses  
  
the work of Thurs-  
idson (17), Porter  
validity of attitude  
estions. It is per-  
1929) was the only  
on the Internation-  
the highest possible  
the highest (anti-)  
nt who has recently  
ivities on the Colum-

DIFFERENT GROUPS ON THE

Internationalism S.D.	Negro	
	M	S.D.
7.34	56.13	6.31
7.10	55.53	6.77
5.95	55.87	5.12
6.98	59.40	6.09
6.56	54.22	6.22
6.98	60.21	5.21
7.11	56.44	6.22
7.04	54.73	5.81
	46.35	7.31

TABLE VIII  
COMPARISON OF THE THURSTONE-DROBA WAR SCALE AND THE  
INTERNATIONALISM SCALE

Group	N	Reliability Coefficients Thurstone-Droba Scale		Reliability Coefficients of International- ism Scale	Internationalism Scale vs. Thurs- tone-Droba Scale	
		Form A vs. B 22 Items vs. 22	Corrected A & B 44 Items	24 Items	Raw	Corrected For At- tenuation
C	54	.78	.88	.88	.71	.81
F	32	.74	.85	.88	.65	.75

In view of the fact that the method presented here when compared with the Thurstone method gave evidence of yielding the same reliability with fewer items, or higher reliabilities with the same number of items, it was decided to try the 1 to 5 method of scoring upon the Thurstone-Droba War scale to see how it would compare with Thurstone's method of scoring. Using Group C each individual was asked to indicate whether he strongly agreed, agreed, was undecided, disagreed, or strongly disagreed with each statement in the Thurstone-Droba War scale, Forms A and B.

Four statements in each form were not used in the scoring because it was found virtually impossible to determine whether to assign a value of 1 or 5 to the "strongly agree" alternative. An illustration of such a statement is number 5 in Form A: "Compulsory military training in all countries should be reduced but not eliminated." It is impossible to tell whether a person is agreeing or disagreeing with the "reduction" aspect of this statement or the "not eliminated" aspect. A person who strongly opposes compulsory military training would disagree or strongly disagree with the "not eliminated" aspect, whereas a person who favors compulsory military training would disagree or strongly disagree with the "reduction" aspect of the statement. Obviously for the 1 to 5 method of scoring the statement is double-barreled and of little value because it does not differentiate persons in terms of their attitudes. Persons at either extreme of the attitude continuum can readily check the same alternative.

Another illustration of a statement that could not be used is number 17 of Form B: "Wars often right tremendous

wrongs." T  
fact, and co  
son regardl  
not used fo

Fc  
Fc

The crit  
50-52 was  
numerical  
each stater  
pected.

The resu  
IX. The 1  
each form  
as the Thu  
most plaus  
by the 1 to

COMPARISON  
THURST

Thurstone-D  
(18 quest  
stead of 2

Regular Th

COEI

Thurstone  
of Thurst

The two  
.83, which  
is possible  
methods,  
ing, a sti  
methods

wrongs." This might be treated as a statement concerning fact, and could well be agreed with or disagreed with by a person regardless of his attitude. The other statements that were not used follow:

- Form A, statements number 8, 10 and 17.
- Form B, statements number 5, 10 and 20.

The criterion of internal consistency, discussed on pages 50-52 was used as an objective check to see (1) whether the numerical values were properly assigned and (2) whether each statement differentiated the extremes in the manner expected.

The results expected were obtained and are shown in Table IX. The 1 to 5 method of scoring with fewer items used on each form yielded as high a reliability coefficient for *one* form as the Thurstone method did for the *two* forms combined. The most plausible explanation for this higher reliability obtained by the 1 to 5 method has already been suggested on page 24.

TABLE IX

COMPARISON OF THE THURSTONE AND THE 1 TO 5 METHOD OF SCORING THE THURSTONE-DROBA WAR SCALE DATA FROM GROUP "C" (N = 54)

	Form A vs. B	
	Raw	Corrected
Thurstone-Droba scale scored 1-5 method (18 questions only used in each form instead of 22)	(18 vs. 18) .88	(36 items) .94
Regular Thurstone scoring	(22 vs. 22) .78	(44 items) .88
COEFFICIENT OF CORRELATION BETWEEN THE TWO METHODS		
	Raw	Corrected for Attenuation
Thurstone scale (44 items) vs. 1-5 scoring of Thurstone scale (36 items)	.83	.92

The two methods of scoring correlate quite highly, namely .83, which when corrected for attenuation becomes .92. It is possible that if the same statements had been used in both methods, rather than four less in each form on the 1 to 5 scoring, a still higher coefficient of correlation between the two methods would have been obtained.

LE AND THE

Internationalism Scale vs. Thurstone-Droba Scale

Raw	Corrected For Attenuation
.71	.81
.65	.75

nted here when  
vidence of yield-  
higher reliabili-  
ecided to try the  
Droba War scale  
method of scor-  
sked to indicate  
ecided, disagreed,  
a the Thurstone-  
ed in the scoring  
le to determine  
"strongly agree"  
nent is number 5  
; in all countries  
is impossible to  
greeing with the  
"not eliminated"  
mpulsory military  
ree with the "not  
favors compulsory  
gly disagree with  
obviously for the 1  
ole-barreled and of  
e persons in terms  
me of the attitude  
native.  
could not be used is  
right tremendous

Read

## V. PSYCHOLOGICAL INTERPRETATION

From the foregoing it will be clear that at least three general or group factors in social attitudes have been discovered within the student populations which participated in our study. The results require a closer scrutiny of the problem of the psychological significance of such group factors, as contrasted with the possibility of their existence as mere statistical artifacts.

From the work of many character testers, notably May and Hartshorne, the impression has gotten abroad that behavior is almost entirely controlled by the details of particular situations, the intercorrelations between different forms of honesty, for example, being exceedingly low. It is shown by the members of the Character Education Inquiry that prediction from scores in one situation to scores in another situation is but little better than a chance shot. A copious literature has already arisen in which the plea is made for indoctrination of children in specific moral habits, as contrasted with the teaching of general moral precepts. That all this discussion is really beside the point will be evident in the light of two facts: first, that the largest number of tests devised by May and Hartshorne for any single character trait was nine, this being the number used in measuring "honesty," while those involved in measuring "service" and "self-control" were even fewer. The nine tests for honesty were deliberately chosen to cover a wide "honesty area." The average inter- $r$  of the entire nine tests was .227. Substituting this figure in the Spearman-Brown formula a predicted reliability of .725 was obtained (11, II p. 125). It will be noticed that this does not really indicate the specificity which is ordinarily assumed to have been proven by the Character Education Inquiry. Furthermore, the reader will note that May and Hartshorne themselves indulge in predictions as to the number of tests which would have to be devised in order to cover the *entire honesty area*, the *entire service area*, and the *entire self-control area*. The assumption that such an "entire area" exists in a psychological and not merely a statistical sense, assumes, of course

that a *general honesty factor exists*. On this basis the only significance of the data on specificity is in showing that the particular moral habits studied are too few and too distantly related to give much prediction from one to the next. Not only could other tests have been chosen which would have intercorrelated higher, but a longer battery of similar tests would have revealed more clearly the magnitude and significance of a general honesty factor. It is in fact obvious that the specificity or generality of a trait, as measured, depends very largely upon the number and variety of situations used in the testing. A trait like honesty is either general or specific according to the number and degree of similarity of one's tests. On the other hand, the Character Education Inquiry showed lower specificity in *honesty attitudes* in paper and pencil situations than in the overt behavior situations. There is, therefore, a reason to suspect that the conversion of behavior dispositions into verbal form tended to make for an increase in internal consistency; but this is an hypothesis of rather limited value, since of course, all such results will depend upon the actual attitude tests used.

What, then, can we mean by affirming that the scaling methods here reported indicate high generality rather than specificity in social attitude; what is there in the present results which is psychologically new?

The reply requires a consideration of the difference between what may be called the biological-science point of view and the social-science point of view, in relation to attitudes. From the point of view of one trained in the biological sciences, an attitude is considered to be a disposition of the organism, a disposition depending, of course, upon the constitution and the whole life history of the individual. For convenience, one may say that an attitude is a conditioned response to certain verbal formulae. Attitudes will cluster together insofar as the verbal stimuli are similar, or insofar as the individual's training has resulted in the establishment of a final common path for several verbal propositions. We should expect, of course, that attitudes toward the use of the marines in Haiti would differ but little from attitudes toward the use of the marines in Nicaragua, because the verbal propositions are parts of almost indistinguishable contexts of verbal situations in which the problem of the use of force to protect investment in Latin-

## ATION

east three gen-  
een discovered  
ipated in our  
the problem of  
actors, as con-  
e as mere sta-

otably May and  
that behavior is  
articular situa-  
arms of honesty,  
wn by the mem-  
prediction from  
situation is but  
s literature has  
ndoctrination of  
l with the teach-  
is discussion is  
ght of two facts:  
sed by May and  
as nine, this be-  
" while those in-  
ontrol" were even  
liberately chosen  
ge inter-r of the  
his figure in the  
bility of .725 was  
that this does not  
arily assumed to  
on Inquiry. Fur-  
Hartshorne them-  
per of tests which  
the *entire honesty*  
: *self-control area*.  
exists in a psycho-  
ssumes, of course,

America is involved. Even propositions regarding the use of force in protecting investments in China will presumably have been affected to a large degree by the same influences.

Furthermore, there exist empirically numerous clusters of attitudes in which the stimulus situations do not appear on the surface to be similar. Why the attitude toward the Jim Crow car and the attitudes toward a Negro's buying a home or farm should be closely linked is not apparent if one considers merely the *direct* social implications of the two. It is only when one realizes that both propositions touch off a general attitude toward the Negro that one sees the reason for the linkage. Quite in contrast with our expectations, we are obliged to report that there exists a clear-cut pro- or anti-Negro sentiment, an emotional and conative disposition which runs through the entire fifteen items used to study white attitudes toward the Negro. Specific factors there indeed are, as may be clearly seen by comparing each item with the next, but even with as few as fifteen items, the general factor of a favorable (or unfavorable) attitude stands out clearly. This, then, is a psychological fact transcending the mere statistical aspects of the scaling method involved.

At this point it will be natural to raise the objection that the high generality of such group factors is entirely a result of the special social conditions surrounding the education of the particular individuals who acted as subjects in our study. This point should be not only conceded but emphasized. While from the biological-science point of view, it may be sufficient to point to the existence of general conative dispositions, it is valuable from the social-science point of view to recognize that these group factors are aspects of the cultural patterns shared by the 650 individuals here tested. In this way, the study of the degree of similarity of results from one university to another offers an empirical method of establishing the degree of uniformity of the cultural pattern in which these various groups of students participate. As pointed out by Rice (28, pp. 125-175), the concept of the cultural area is capable of quantitative definition; and what may be called college student attitudes in the northeastern United States toward the Negro are shown quantitatively to be amazingly alike in the following points:

In response to the question: "In a community in which the

negroes outnumber the whites, under what circumstances is the lynching of a negro justifiable?", all of our northern college student groups chose the response "Never" far more often than any of the other alternatives; in fact 65 to 75 per cent of each of our northern groups gives this response. Clear-cut "public opinion" on this issue appears. On the other hand, among 123 students at an institution in Virginia only 31 per cent chose this answer. The difference between "general" northern and southern attitudes toward the Negro appears in striking fashion in Table VII, being, as has been noted, many times as large as it would need to be in order to be reliable statistically. In these cultural areas points of great *difference* will nevertheless be seen.

A scaling method such as that employed here seems to have value as a way of revealing general differences in the friendliness or antagonism shown by different social groups toward some nationality or class with which they have dealings. Not only is the item-by-item analysis of attitudes of different college groups of some interest, but the tendency of one college group to differ significantly in its general favorableness or unfavorableness in attitude on war, internationalism, and the Negro may be brought into line with such facts as the following: the geographic location of the college; the social and economic status of the students' parents, the influences of preparatory school and the various courses taken by the student in college, the reading habits, extra-curricular activities and other personal items upon which data are available. It is not the purpose of the present paper to survey systematically the reasons for some of the major differences appearing among the various college groups compared. Our interest lies rather in showing that even these short and simple scales do show clear-cut and reliable differences in attitude, and that by their very nature they bring out such differences more significantly than could any scaling method with lower reliability or any method in which specific, rather than general, attitude factors were emphasized. If the present method has any value at all, it has value as a device for laying bare the general favorableness or unfavorableness which certain individuals and groups display toward other individuals and groups. To what extent the cultural patterns reflecting themselves in the students' attitudes are "constant" and homogeneous will be evident

from the study of the reliability of the scales themselves and from a study of the range and sigmas of the scores. The answer must, of course, be that within each college group very large variations are found; for example, though in the northern college group some very strongly anti-Negro papers and in the south some very strongly pro-Negro papers are found, it will nevertheless be seen that sharply contrasting general trends do exist. The factors which make for favorableness or unfavorableness are legion, but the relative number and strength of favorable forces is vastly greater at some colleges than at others.

The degree of interrelationship between our three scales has already been indicated. It will be recalled that the uncorrected correlation between the Negro and Internationalism scales is  $+.40$ ; that between Negro and Imperialism,  $+.34$ ; that between Imperialism and Internationalism,  $+.63$ . Material from three other universities yielded very similar results. As would be expected, the common ground between anti-imperialism and internationalism is higher than that which appears in other comparisons. The question will naturally arise here whether some general characteristic of the student tends to make for high scores in all three fields, and thus to produce perhaps a spurious correlation between them. The question cannot be definitely answered, but in view of the frequent argument that general intelligence is responsible for such intercorrelations as these, it is worth noting that a correlation was sought in Group C between internationalism and intelligence (Otis Self-Administering Higher Examination, Form A, 55 cases) and turned out to be  $-.17$ ; in Group F the correlation between the same two scores is  $+.01$  (33 cases). In view of these slight correlations, the point was not judged worth following further in the present study. Cultural causes rather than innate differences seem likely to be responsible for the interrelations found.

Having insisted on the social or cultural character of the determining forces involved, not only in specific attitudes but in general attitude trends, it may be of interest to note that some very suggestive evidence exists here as to the presence of even more highly generalized attitude factors which at present cannot be proven to be cultural. We have already noted the correlation of  $+.40 \pm .06$  between pro-Negro and pro-inter-

nationalism scores, and  $+.34 \pm .06$  between pro-Negro and anti-imperialism scores. These correlations, although not very high, are reliably greater than zero, and indicate according to the logic used above a significant general factor running through these three scales. This, we feel compelled to admit, looks suspiciously like the general radicalism or conservatism factor which we were at first disposed to discount. While insisting that this most general factor is not very large as compared with the group factors involved in the three scales, it must, nevertheless, be conceded that there seems to be a tendency for the protest against existing social institutions and practices to exist as a definite trait reappearing throughout our series of questions. This is supported by a preliminary analysis of some of the personal data on our subjects, which leads us to suspect, at this writing, the presence of what one might call contented and discontented attitude factors, so broad as to be practically equivalent to the aforementioned radicalism and conservatism factors. It is possible that such general factors as radicalism or conservatism are due to the innate restlessness, suggestibility, etc. of the persons tested, or on the other hand to the homogeneity of certain opinion-making forces in society, the fact, for example, that journals which reflect radical opinion regarding Negro-white relations usually reflect at the same time radical international and economic opinions. We have no intention of pleading here either way as regards the point, but since we have affirmed with some degree of confidence the primarily cultural character of the group factors measured in this study, we think it appropriate to add that in our opinion the explanation of the general radicalism and conservatism factors is an open question, as likely to be settled one way as the other.

## VI. SUMMARY

X This study has presented a different method of measuring attitudes from that ordinarily used. It has attempted to evaluate this method, both in its longer, more complex form and in its simpler form, and to compare it with other methods widely used. The results obtained seem to justify the following statements:

1. Assuming that attitudes are distributed normally, a method of measuring attitudes has been developed which uses sigma units. This method not only retains most of the advantages present in methods now used, such as yielding scores the units of which are equal throughout the entire range, but it has additional advantages. These briefly are: first, the method does away with the use of raters or judges and the errors arising therefrom; second, it is less laborious to construct an attitude scale by this method; and third, the method yields the same reliability with fewer items.

14/15 18 11 X  
2. A simpler method was found which yielded essentially the same reliability coefficient as the sigma method and correlated so well with this (+.99) that it seems justifiable for all ordinary purposes to use the simpler method. The simpler technique involved the assigning of consecutive numerical values to the different alternatives. (On pages 15-20 the numbers in parentheses below the different alternatives show the numerical values assigned.) After assigning the numerical values to the different possible responses, the score for each individual was determined by finding the average or sum of the numerical values of the alternatives that he checked.

3. A scale measuring attitudes toward international relations constructed by the simpler method when compared with the Thurstone-Droba War scale yielded the same reliability coefficients with practically one-half as many items. The coefficients of correlation between the two scales for two different groups was respectively +.71 and +.65. When corrected for attenuation these coefficients became +.81 and +.75. This relationship is obtained in spite of the fact that the statements in the two scales are quite dissimilar.

4  
mer  
twe  
coef  
Thu  
fort  
the  
tion  
cor  
5  
in a  
stru  
6  
tecl  
in p  
pec  
alis  
spe  
que

4. Applying this simpler method of scoring to the statements in the Thurstone-Droba scale and using eighteen of the twenty-two statements in each form the uncorrected reliability coefficient obtained for Form A versus Form B was .88. The Thurstone method of scoring for all twenty-two items in each form yielded an uncorrected coefficient of correlation between the two forms of .78. The uncorrected coefficient of correlation between the scores obtained by the two methods was  $+.83$ , corrected,  $+.92$ .

5. Methods of objectively checking the statements included in a scale are presented to assist the experimenter in the construction of attitude scales.

6. While the present monograph aims chiefly to describe a technique rather than generalized results, it is noteworthy that in practice the three scales yield, contrary to the writer's expectation, very clear-cut generalized attitudes on "internationalism," "imperialism," and "the Negro," in addition to the specific attitudes which appear in response to the separate questions.

of measuring  
 ted to evalu-  
 form and in  
 ethods widely  
 llowing state-

normally, a  
 ed which uses  
 ost of the ad-  
 yielding scores  
 ire range, but  
 are: first, the  
 udges and the  
 orious to con-  
 rd, the method

lded essentially  
 ethod and cor-  
 s justifiable for  
 d. The simpler  
 ative numerical  
 ages 15-20 the  
 lternatives show  
 ing the numeri-  
 he score for each  
 erage or sum of  
 he checked.

international rela-  
 n compared with  
 ame reliability co-  
 items. The coef-  
 s for two different  
 hen corrected for  
 and  $+.75$ . This  
 hat the statements

## APPENDIX

### THE METHOD OF CONSTRUCTING AN ATTITUDE SCALE

#### 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 importance 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*. The use of *should* is well illustrated in the "strongly approve" type of statements shown on pages 15-20. X

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 statement 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 each statement so worded that the modal reaction to it is approximately in the middle of the possible responses. X

X 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.

X 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:

Numerical Value

1 "How far in our educational system (aside from trade education) should the most intelligent negroes be allowed to go?  
 (a) Grade school.  
 (b) Junior high school.  
 (c) High school.  
 (d) College.  
 (e) Graduate and professional schools."

2  
 3  
 4  
 5

5 "In a community where the negroes outnumber the whites, a negro who is insolent to a white man should be:  
 (a) excused or ignored.  
 (b) reprimanded.  
 (c) fined and jailed.  
 (d) not only fined and jailed, but also given corporal punishment (whipping, etc.).  
 (e) lynched."

4  
 3  
 2  
 1

"All negroes belong in one class and should be treated in about the same way."

Value

STRONGLY APPROVE	APPROVE	UNDECIDED	DISAPPROVE	STRONGLY DISAPPROVE
(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."

Value

STRONGLY APPROVE	APPROVE	UNDECIDED	DISAPPROVE	STRONGLY DISAPPROVE
(5)	(4)	(3)	(2)	(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. Thus, if one wishes to call favorable to the Negro the following responses to the statements shown on pages 18-19, no serious objections will be raised providing that he is consistent in his designations; to do so, however, does seem to be less reasonable and not in accord with common usage.

Statement Number	Response
1	Yes
2	Yes
3	No
4	No
5	No
6	(e)
7	(a)
8	(e)
9	Strongly approve
10	Strongly approve
11	Strongly disapprove
12	Strongly disapprove
13	Strongly approve
14	Strongly disapprove
15	Strongly approve

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 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 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 psychologically 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 re-

quired 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 X shows a comparison of the results obtained from item analysis and the criterion of internal consistency. 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.

TABLE X

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

rho (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.

In using the criterion of internal consistency the reactions of the group that constitute one extreme in the particular atti-

tude 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 XI 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

of at-  
noted  
calcula-  
abori-  
nd the  
results  
the re-  
nternal  
en the  
rmined  
ency as  
nal con-  
et yields

N OF THE  
HE NEGRO  
)

Column 5

5
6
11
14
8
13
15
10
12
1
9
7
2
3
4

he individu  
tatements.  
est 9 indivi

alysis bas  
olumn 2.  
iteration of  
es shown

he reaction  
rticular at

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.

TABLE XI

CRITERION OF INTERNAL CONSISTENCY APPLIED TO THE INTERNATIONALISM SCALE FOR GROUP "D"—(N = 100)

HIGH GROUP		STATEMENT NUMBERS																							
		<i>Three-Point Statements</i>												<i>Five-Point Statements</i>											
<i>Indiv. No.</i>	<i>Score</i>	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	3	5	4
13	102	4	4	4	4	4	4	4	4	4	4	4	4	5	5	3	5	5	5	5	5	5	3	5	5
10	101	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	4	5	5	5	5	5	4	5	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	5	5	3	5	5	5	3	4
27	98	4	2	4	4	4	4	4	4	4	4	4	4	4	5	3	5	5	5	4	4	5	2	5	5
60	98	4	4	4	4	4	4	2	4	2	4	4	4	5	5	3	4	5	5	3	5	5	4	5	5
64	98	4	4	4	4	4	4	4	4	2	4	4	4	4	3	4	3	5	5	5	5	5	4	5	4
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	35	41	42
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	17	14	22
D		16	14	15	8	12	7	11	16	10	15	2	13	21	13	13	25	14	23	25	22	21	18	27	20
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	2	1	2
77	54	2	2	2	2	2	4	2	2	2	3	4	2	2	3	3	1	4	2	1	2	2	2	1	2
22	60	2	2	2	2	4	4	2	2	2	2	4	2	4	3	1	3	4	2	2	2	4	2	1	2
35	61	2	2	2	4	3	3	2	2	4	4	4	3	2	2	3	1	3	2	3	1	3	2	2	2
53	62	2	2	2	4	4	2	2	2	2	2	4	2	2	3	2	2	2	3	1	3	1	1	1	1
69	62	2	2	4	2	2	4	2	2	2	2	4	2	2	3	3	2	4	1	1	4	5	2	1	4
94	63	2	2	2	4	3	4	2	2	4	2	4	4	2	2	3	2	4	2	2	3	2	2	2	2
21	64	2	2	2	4	2	4	3	2	2	2	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	2	4	4	4	3	3	1	4	4	2	2	1	2	3	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	17	14	22

MEASUREMENT OF ATTITUDES

## BIBLIOGRAPHY

1. Allport, F. H. and Hartman, D. A. The measurement and motivation of atypical opinion in a certain group. *Am. Pol. Sci. Rev.*, 1925, 19, 735-760.
2. Allport, G. W. The composition of political attitudes. *Amer. J. Sociol.*, 1929, 35, 220-238.
3. Allport, G. W. and Vernon, P. E. The field of personality. *Psychol. Bull.*, 1930, 27, 677-730.
4. Bain, R. Theory and measurement of attitudes and opinion. *Psychol. Bull.*, 1930, 27, 357-379.
5. Chave, E. J. and Thurstone, L. L. The measurement of social attitudes. *Attitude Toward God, Scale No. 22*. Chicago. Univ. of Chicago Press. 1931.
6. Droba, D. D. The measurement of social attitudes. *Attitude Toward War*. Chicago. Univ. of Chicago Press. 1930.
7. Filter, R. O. An experimental study of character traits. *J. Appl. Psychol.*, 1921, 5, 297-317.
8. Folsom, J. K. *Social psychology*. New York. Harpers. 1931.
9. George, R. W. A comparison of Pressey X-O scores with liberal-conservative attitudes. Master's essay in Columbia Univ. Libr. 1925.
10. Hartmann, G. W. Precision and accuracy. *Arch. Psychol.* 1928, No. 100.
11. Hartshorne, H. and May, M. A. *Studies in deceit*. New York. Macmillan. 1928.
12. Hartshorne, H., May, M. A. and Maller, J. B. *Studies in service and self-control*. New York. Macmillan. 1929.
13. Hartshorne, H., May, M. A. and Shuttleworth, F. K. *Studies in the organization of character*. New York. Macmillan. 1930.
14. Hinckley, E. D. A scale for measuring attitude toward the negro. Chicago. Univ. of Chicago Press. 1930.
15. Hunter, C. W. A comparative study of the relationship existing between the white race and the negro race in the State of North Carolina and in the City of New York. Master's essay in Columbia Univ. Libr. 1927.
16. Katz, D., Allport, F. H. and Jenness, M. B. Students' attitudes; a report of the Syracuse University Reaction Study. Syracuse. Craftsman Press. 1931.
17. Kulp, D. H., II, and Davidson, H. B. Can Neumann's "Attitude Indicator" be used as a test? *Teach. Coll. Rec.*, 1931, 32, 332-337.
18. Maller, J. B. *Character and personality tests*. New York. Teach. Coll. 1932.
19. Mathews, C. O. The effect of the order of printed response words on an interest questionnaire. *J. Educ. Psychol.*, 1929, 20, 128-134.
20. Moore, H. T. Innate factors in radicalism and conservatism. *J. Abn. and Soc. Psychol.*, 1925, 20, 234-244.
21. Murphy, G. *An historical introduction to modern psychology*. New York. Harcourt, Brace and Company. 1929.
22. Murphy, G. and Murphy, L. B. *Experimental social psychology*. New York. Harper. 1931.
23. Neumann, G. B. A study of international attitudes of high school students. *Teach. Coll. Contrib. Educ.* 1927. No. 239.
24. Newcomb, T. M. The consistency of certain extrovert-introvert behavior patterns in 51 problem boys. *Teach. Coll. Contrib. Educ.* 1929. No. 382.

25. Porter, E. Student opinion on war. Doctoral dissertation in Univ. of Chicago Libr. 1926.
26. Rice, S. A. Report, Inst. of methods of rural sociol. research. U. S. Dept. of Agriculture. 1930. 11-20.
27. Rice, S. A. Statistical studies of social attitudes and public opinion. In Rice, S. A. (Ed.), *Statistics in social studies*, pp. 171-192. Philadelphia. Univ. of Pennsylvania Press. 1930.
28. Rice, S. A. *Quantitative methods in politics*. New York. Knopf. 1928.
29. Stouffer, S. A. An experimental comparison of statistical and case history methods of attitude research. Doctoral dissertation in Univ. of Chicago Libr. 1930.
30. Thorndike, E. L. An introduction to the theory of mental and social measurements. Second Edition. New York. Teach. Coll. 1913.
31. Thurstone, L. L. A law of comparative judgment. *Psychol. Rev.*, 1927, *34*, 273-286.
32. ———. An experimental study of nationality preferences. *J. Gen. Psychol.*, 1928, *1*, 405-425.
33. ———. A scale for measuring attitude toward the movies. *J. Educ. Res.*, 1930, *22*, 89-94.
34. ———. Attitudes can be measured. *Amer. J. Sociol.*, 1928, *33*, 529-554.
35. ———. Commentary. In Rice, S. A. (Ed.), *Statistics in social attitudes*, pp. 192-196. Philadelphia. Univ. of Pennsylvania Press. 1930.
36. ———. The measurement of opinion. *J. Abn. and Soc. Psychol.*, 1928, *22*, 415-430.
37. ———. Theory of attitude measurement. *Psychol. Rev.*, 1929, *36*, 222-241.
38. ——— and Chave, E. J. *The measurement of attitude*. Chicago. Univ. of Chicago Press. 1929.
39. Trow, W. C. The psychology of confidence. *Arch. Psychol.*, 1923, No. 67.
40. ———. Trait consistency and speed of decision. *School and Soc.* 1925, *21*, 538-542.
41. Vetter, G. B. The measurement of social and political attitudes and the related personality factors. *J. Abn. and Soc. Psychol.*, 1930, *25*, 149-189.
42. Wang, C. K. A. and Thurstone, L. L. The measurement of social attitudes. Scale No. 21, Forms A, B. *Attitude toward birth control*. Chicago. Univ. of Chicago Press. 1930.
43. Watson, G. Measures of character and personality. *Psychol. Bull.*, 1932, *29*, 147-176.