

EMPIRICAL ASSESSMENT OF ATTITUDES

John W & Deborah J Bardo, Wichita State University
Steven R Steiber, American Medical Association

INTRODUCTION Since Thurstone's landmark article, "Attitudes can be measured" (1928), sociologists have used attitude research extensively. There are now thousands of published papers purporting to measure attitudes, and using attitudinal variables. They vary widely in complexity and sophistication of method. Measurement techniques range from simple response to a single question to multidimensional and factor scaling. Samples may be drawn from only one undergraduate college class, or may involve 15 or more samples from as many countries. From a review of attitude measurements published from 1971 through 1975 in 8 major sociology journals we find that 1) measurements are not necessarily related to any attitude theory; 2) little attention is paid to possible parameters of the construct being measured; 3) reliability is not reported; and 4) item analyses are not reported.

. Three key concepts in attitude research are: attitude, reliability and validity. Attitude is the construct being measured, and reliability and validity concern the adequacy of measurement. For an attitude measure to have any scientific value, it must reliably and validly measure the respondent's disposition toward a social object. Unreliable or invalid measures, despite mathematical rigor, are useless and misleading. The major fault with attitude analysis is that sociologists all too often ignore these factors, that potentially invalidate the study. A review of 841 attitude related articles shows that most of them inadequately report reliability and validity.

RELIABILITY & VALIDITY

. In the social sciences, the term reliability applies to the principle of analytical consistency (Lundberg 1942; Goode & Hatt 1952; Festinger 1953; Bohrnstedt 1970). A reliable research method

must provide consistent qualitative or quantitative observations when applied to a group of subjects under specific conditions. With a technique of known reliability, variations in observations are due mainly to real differences among the subjects, changes in the conditions of observation, or actual modification in the condition being observed. An instrument is reliable when it consistently produces the same results when applied to the same sample under the same conditions. To determine consistency of measure, one either compares two or more time-sequenced measures, as in test-retest consistency, or one analyzes a single set of cross-sectional data for regularity in discrimination among respondents with differing levels of affect, or internal consistency. It should be possible with a reliable instrument consistently to rank individuals according to their measured attitude - the basic principle of attitude measurement.

. Fundamental to the concept of reliability is that: 1) consistency cannot be assumed a priori - it must be demonstrated; 2) measurement consistency is data-bound. Even if an instrument measures consistently in one group, it cannot be assumed that it will do so in any other.

. Major attitude theorists agree that reliability is a data-bound and a time-bound construct. "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 on which it is used. It is certainly reasonable to suppose that just as an intelligence test which has been standardized on one cultural group is not applicable to another, so an attitude scale which has been made for one cultural group will hardly

be applicable in its existing form to other cultural groups" (Likert 1932 53). According to Guttman (1950), a universe of items may form a scale for a population at one time and not at another, because of changes in the meaning of the attitude referent, as with changes of kind rather than degree. A series of items is a scale only if it approximates unidimensionality with a coefficient of reproducibility of .90 or more. Minimum consistency requirements for Likert scales vary, but generally as set of items for which a Spearman-Brown coefficient or Cronbach's alpha of at least .7 is obtained is reported as scaled. Conversely, a universe of items may scale at one time, but not at another, which would show that the attitude had undergone structuralization. A universe may form a scale in one population, and not in another, or the same items may form two different scales in two populations. And a universe may form a scale for a sub-population and not for the entire population. Here the attitude referent has different meaning for various subgroups of the population.

. Both Guttman and Likert argue that reliability of an instrument is specific to the group on which it is determined. It is not exportable, cannot be assumed, and should be reassessed for each application of an instrument.

. Validity ". indicates the degree to which an instrument measures the construct under investigation" (Bohrnstedt 1970 91). For an attitude measure to be valid, it must actually measure its referent. Validity is indicated by logical or face validity, construct validity, content validity, jury or known-groups validation, and criterion-related validity (Goode & Hatt 1952 237). Like reliability, validity is also data-bound. An instrument may validly measure one population, and not another, a single population on one occasion, but not on another; or for a subgroup, but not the rest of the population. Validity should be demonstrated for each application

of an instrument.

. Reliability and validity are distinct but related constructs. An attitude measure may be reliable but not valid; however, validity does imply reliability. Unless observations are made regularly, one cannot argue that any particular construct is truly measured. The more consistently an instrument measures, the more accurate, and thereby the more valid it may be. Reliability is a necessary, but not a sufficient condition for validity. One cannot convincingly argue that a social construct is measured unless the instrument is demonstrably reliable and valid. And the instrument should be retested each time it is applied to a new population, or under different conditions, or after a significant period of time.

. These arguments are elementary, and one would assume that the well-trained social scientist would know the stipulations of reliable and valid measurement. Empirical research on Thurstone, Likert, and factor scaling reinforce the warnings about data-boundness in attitude measures (Flynn & Carter 1972; Bardo 1976; Kinder & Reeder 1975). This highlights the fact that many studies are published in which reliability and validity are unassessed.

ATTITUDES Attitudes are conceived in social psychology as emotion-based dispositions to some referent external to the person (Thurstone 1928; Allport 1935; Katz & Stotland 1959; Katz 1960; Kidder & Campbell 1970). Summers (1970) sees substantial agreement among attitude theorists on 4 points. Attitudes: 1) consist of predispositions to respond, rather than to act; 2) persist over time; 3) (should) produce consistent behavior; and 4) have directional or motivating qualities.

. Attitudes are not observable, and hence are not directly measurable. Their existence, intensity, and level must be inferred from indirect indicators, usually called scales. Scales can take many forms, and the measurement

and development of indicators of attitudes is a difficult task (Cook & Sellitz 1964; Festinger 1953).

. Using scales without checking reliability and validity can affect the value of any study. 1) When an attitude measure is a basic element of an empirical analysis, and is of unknown reliability and validity, any analysis based on the measure is doubtful. 2) Since unreliability is a form of measurement error, failure to retest and revise instruments brings loss of explanatory power. 3) Unreliable measures can lead to artificially reduced predictability. We have defined attitudes as dispositions to respond. A major issue in attitude research is the attitude-behavior consistency controversy. Attitudes are not seen as good predictors of behavior, because too many other conditions intervene. Since most attitude measures are of unknown validity, a part of their inability to predict is due to inefficient measurement.

METHOD The literature review included attitude related articles published between 1971 and 1975 in 8 journals: American Sociological Review, American Journal of Sociology, Social Forces, Sociological Quarterly, Sociology & Social Research, Sociometry (Social Psychology Quarterly), Public Opinion Quarterly, and Journal of Social Psychology. Selections were based on attitude-related content of the abstracts and summaries.

FINDINGS The data revolve on 4 questions: 1) how are attitudes conceived; 2) how are they measured; 3) how is reliability determined; and 4) what are the major sources of attitude instruments?

Conceptions of attitudes. The theoretical debate over the complex nature of attitudes is not reflected in empirical attitude studies. In two-thirds of the articles theoretical rationale for instrument choice is limited or missing. Researchers define the attitude as what the scale measures, for Bogardus' social distance scale, Cattell's personality scale, & Srole's

anomia scale. They assume that if a preconstructed scale is used, the originator determined the relevant universe of items, and that further justification is needless. If the scale is newly developed for the study, it is assumed that it adequately taps all facets of the attitude in the population. These assumptions seem to hold even when a single item is the attitude indicator.

. While it is likely that many researchers seriously consider the nature of the attitude referent, in the empirical literature, little attention is paid to the possible scope, domain, or other parameters of the attitude referent. The attitude is defined most often as what the scale measures.

How attitudes are measured. The most popular form of attitude measure is the Likert scale. Over 40 percent of all instruments used it. Next was the forced choice format (20%), followed by semantic differentials, Guttman scales, and ranking and rating scales (about 6% each). Thurstone scales are rarely used and account for 1 percent of attitude instruments.

. What measurement means also varies. In 16 percent of interpretable cases, attitude measure involved no more than asking subjects to respond to a single statement or series of separately analyzed statements. To use single-item indicators becomes a critical problem when viewed in the light of attitude theory. The researcher must assume that all relevant facets of the attitude referent, however complex, can be reliably summarized in that single item.

Determination of reliability. In practice, this is one of the easiest tasks, yet the most neglected. Of more than 800 measurements, only 30 percent attempted to measure and report reliability. The number of reliability reports are indicated by journal in Table 1. There were differences among the journals, but less than half of the instruments included any consideration of reliability. A second question concerns changes over time. The proportion of attitude

TABLE 1: REPORTED RELIABILITY ON ATTITUDE INSTRUMENTS

Journal	Articles	Scales	Reported Reliability
American Sociological Review	51	138	33%
American Journal of Sociology	46	86	36
Social Forces	36	39	44
Sociological Quarterly	30	52	21
Sociology and Social Research	40	44	21
Sociometry	55	81	35
Journal of Social Psychology	187	350	23
Public Opinion Quarterly	34	51	28

TABLE 2: RELIABILITY REPORTED BY YEAR

Year	N	Reported Reliability
1971	160	22%
1972	175	31
1973	162	36
1974	111	20
1975	233	28

TABLE 3: REPORTED RELIABILITY BY SAMPLE TYPE

Type Sample	N	Reported Reliability
Polyethnic	72	24%
Cross-cultural	65	25
Non-American	159	28

measurements with reported reliability is roughly stable in the years studied. There was no tendency for improvement over time, as shown in Table 2.

. A third problem in reliability estimates appeared with cross cultural samples. It was reported for about 25 percent of all studies, and 28 percent of the non-American samples, as shown in Table 3. When cultural boundaries are crossed there is even less attention to instrument adequacy.

Methods to determine reliability.

Internal consistency was measured by Cronbach's alpha (30%), split-half correlation (20%), coefficient of reproducibility (18%), Kuder-Richardson item/total correlation (10%), and test-retest statistics (16%). For item analysis, the most used was factor analysis with orthogonal rotation (33%). Item total

correlation was calculated with Cornell or Goodenough methods (12%), and split-half correlation (10%) for 260 cases.

Prebuilt & original instruments.

The prebuilt scale is defined as an attitude measurement scale made and previously used on some other sample. The researcher either gave credit to another social scientist for the scale, or indicated that the scale was originally created for another sample. About 40 percent of the attitude measures were prebuilt scales, and in these cases, only 20 percent were tested for reliability. Usually, researchers using these scales did not discuss reliability or validity, or they referred the reader to other publications, or simply noted the reliability coefficients in the prior application.

DISCUSSION The Willers have created a three-fold typology of knowledge based on specific types of logical connection: 1) empiricist, 2) rational, and 3) abstractive. They criticize sociology for engaging in pseudoscientific systematic empiricism, and that it is committed to becoming an empirical science which is "... often defined as an emphasis on data collection and analysis" (Willer & Willer 1973 1). They contend that sociologists' methods fail to meet the requirements of good empiricism. Empiricist logic requires connection by observation. If logical connections are drawn on a theoretical level, they are rational, and if determined by both theory and observation, the results are abstractive.

. Sociologists' dominant epistemological stance is abstractive. Observation should lead to theoretical generalization which should suggest more observation. In terms of attitudinal measurement, the Willers' arguments are well founded. In the primary sociology journals quoted to our students most attitudinal measures are based on atheoretic instruments of unknown empirical reliability and validity. The value of our abstractive theoretical generalizations is therefore questionable.

. As an empirical science, sociology requires that we base observations on the best available instrumentation. The best available method is subject to difference of opinion, and varies according to circumstances. Regardless of the method, we consider consistency and validity of measurement to be of utmost importance.

. In this research we gathered opinions on cultural bias in attitude measurement from more than a dozen methodologists, some of whom were referees for the journals reviewed in this article. Their general conclusion was: "everybody knows this." We have had papers rejected from several journals and from one meeting on these very grounds. Yet in the same journals for which these individuals referee, most attitude measures published did not report on reliability of instruments. This could have arisen from two factors: 1) reliability was calculated but not reported, or 2) the knowledge considered common by methodologists is not shared by others in the field. Given the propensity of sociology journals to publish reports "sophisticated" in mathematics and intricate in method, we doubt that reliability was calculated but not reported.

. Sociologists involved in method-theory development may in fact understand the reconstructed logical requirements, but in terms of logic in use, that knowledge is not applied (Kaplan 1964). And the fact that this knowledge is not shared by others in the discipline has apparently escaped

the methodo-logicians. Sociology cannot evolve as a sophisticated empirical science until its fundamental logic is manifest in research. We expend much effort creating intricate mathematical models, and are rewarded for it. But the over-riding question remains. How can these models improve our science if its basic logic remains confined to our collective unconscious?

. Most sociologists who measure attitudes use them as elements in broader studies. The attitude is of less interest than that which it is used to predict. Using a scale without reliability and validity tests endangers such a study because: 1) validity is questionable, 2) measurement error may be needlessly high, and 3) predictability may be artificially reduced. Given the high efficiency computing facilities and library programs available to most sociologists, such problems need not continue.

. This research highlights the necessity to reintegrate attitude theory and practice. Attitudes are complex consequents of the individual's socialization process. Attitudes are only indirectly measurable by indicators. Willingness to express attitudes is constrained by a host of biographic and historic factors. Given these conditions, can a single scale measure attitudes in more than one population or more than one condition? For specific attitudinal constructs the answer may be "yes," but it certainly should not be assumed a priori.

REFERENCES

- .Allport Gordon 1935 Handbook of Social Psychology. Worchester Clark U Press
- .Bardo John W 1976 Internal consistency and reliability in Likert-type attitude scales. Sociology & Social Research 60 403-420 July
- .Bohrnstedt George W 1970 Reliability & validity assessment in attitude measurement. Attitude Measurement. G Summers ed. Chicago Rand McNally
- .Cook S, C Selltiz 1964 Multiple

(Concluded on page 180)

same as for the control group with no previous sociology courses whatever. This decay, coupled with the second semester social problems course may partially explain the apparent finding that a second semester of sociology has no effect on student understanding of basic sociological concepts.

The choices of students on the problematic scenarios indicates that an exposure to sociology is positively related to research oriented problem solving for the scenarios. The overall results point to the learning and operationalizing of sociological concepts by our introductory students. This is more heartening that the statement of one student who said: "The most important thing I learned in sociology was that in German, a W is pronounced as a V."

REFERENCES

- .Lenski Gerhard, Richard Lenski 1970 Introduction to Sociology. New York McGraw Hill
- .Light Donald, Suzanne Keller 1979 Sociology. 2nd ed. New York Knopf
- .Popenoe David 1980 Sociology. 4th ed. Englewood Cliffs NJ Prentice Hall
- indicator approach to attitude measurement. Psychological Bltn 62, American Psychological Assn
- .Festinger L, D Katz 1953 Research Methods & the Behavioral Sciences. New York Dryden
- .Flynn C, L Carter 1972 Retest of a measurement model. Sociological Methodology 1972, H Costner ed. San Francisco Jossey Bass
- .Goode W, P Hatt 1952 Methods in Social Research. NY McGraw Hill
- .Guttman Louis 1950 Basis for scalogram analysis. Studies in Social Psychology of World War II, S Stouffer et al, Vol 4 60-90 Princeton U Press
- .Kaplan Abraham 1964 Conduct of Inquiry: Methodology for Behavioral Science. Scranton Penn Chandler
- .Katz Daniel 1960 Functional approach to study of attitudes. Public Opinion Qtrly 24 163-204
- .Katz D, E Stotland 1959 Preliminary statement to theory of attitude, structure, and change. Psychology: A Study of a Science. S Koch ed. NY McGraw Hill 423-475
- .Kidder L, D Campbell 1970 Indirect testing of social attitudes. Attitude Measurement. G Summers ed. Chicago Rand McNally
- .Kinder D, L Reeder 1975 Ethnic differences in beliefs about control. Sociometry 38 216-272
- .Likert Rensis 1932 Technique for measurement of attitudes. Archives of Psychology #140 New York Columbia University
- .Lundburg G A 1942 Social Research: A Study in Methods of Gathering Data. New York Longmans Green
- .Summers G 1970 Attitude Measurement. Chicago Rand McNally 1-17
- .Thurstone L 1928 Attitudes can be measured. American J of Sociology 33 529-554
- .Thurstone L, E Chave 1929 Measurement of Attitude. U of Chicago Press
- .Willer David, Judith Willer 1973 Systematic Empiricism: Critique of a Pseudoscience. Englewood Cliffs NJ Prentice Hall