MEASUREMENT INSTRUMENTS

MURRAY A. STRAUS

BARBARA WAUCHOPE

Standardized instruments for measuring key concepts are vital tools for sociological research. As in other branches of science, new instruments tend to produce a flowering of research focused on the newly measurable concept. The Social Distance Scale, first published in 1928 by Bogardus, illustrates that process. Since then a number of excellent instruments have appeared. Many of them are abstracted or reproduced in compendia such as Sociological Measurement (Bonjean, Hill, and McEmore 1967), Handbook of Research Design and Social Measurement (Miller 1991), Measures of Occupational Attitudes and Occupational Characteristics (Robinson, Athanasiou, and Head 1969), Women and Women's Issues (Beere 1979), Handbook of Scales for Research in Crime and Delinquency (Brodsky and Smitherman 1983), and Handbook of Family Measurement Techniques (Touliatos, Perlmutter, and Straus 1990). The current state of instrumentation in sociology, however, is not as robust as these titles might suggest. The data presented below indicate that the development of new measurement instruments has not kept pace with the growth in sociological research. Moreover, the reliability and validity of most of the instruments have not been established. Explanations for these problems will be suggested.

DEFINITION

Measurement instrument as used in this article is synonymous with terms such as scale (including Likert, Thurstone, Guttman, and Semantic Differential scales), index, test, factor score, scoring system (when used as the indicators for indexes measuring social interaction variables; see Bales 1950), and latent variables constructed by use of a structural equation modeling program. The defining feature of each of these types of instruments is that each is "a measure which combines the values of several variables or items [also called indicators, observations, events, questions] into a composite measure . . . used to predict or gauge some underlying continuum which can only be partially measured by any single item or variable" (Nie et al. 1978, p. 529).

ADVANTAGES OF MULTIPLE-INDICATOR INSTRUMENTS

Multiple-indicator measures are emphasized because they are more likely to be valid than single-indicator measures. It is true that a single good item may be enough and thirty bad ones are useless. However, there are reasons why multiple-indicator measures are more likely to be valid. One reason is that most phenomena of interest to sociology have multiple facets that can be adequately represented only by use of multiple indicators. A single question, for example, is unlikely to adequately represent the multiple facets of "community allegiance." At the macrosociological level, a single demographic characteristic, such as the rate of geographic mobility, is not likely to adequately represent the several aspects involved in "community integration."

A second reason for greater confidence in multiple-indicator measurement instruments occurs because of the inevitable risk of error in
selecting indicators. If a single indicator is used and there is a conceptual error in formulating the indicator or in scoring it, hypotheses that are tested by using that measure will not be supported even if they are true. However, when a multiple-indicator instrument is used, Strauss and Baron (1990) found the adverse effect is limited to a relatively small reduction in validity. In a fifteen-item scale, for example, a defective indicator is only 6.6 percent of the total and therefore will yield findings that are similar to those which would be obtained if all fifteen items were correct. For these and other reasons, a suggested rule of thumb is that three indicators are the minimum needed to reduce the risk involved in staking everything on a single indicator.

Multiple indicators are also desirable because the internal consistency reliability of an instrument as measured by the alpha coefficient (Cronbach 1970) is a function of the number of indicators in the measure and the correlation between them. The more indicators and the higher their correlation with each other, the higher the alpha. If fewer than three items are used, it is almost impossible to achieve a high level of reliability. Reliability needs to be as high as possible because it sets an upper limit on validity. A validity coefficient cannot be greater than the square root of the reliability (Cronbach 1970).

CURRENT STATUS AND TRENDS IN SOCIOLOGICAL MEASUREMENT

There is a rich literature by sociologists on the theory and the procedures needed to construct measures, yet relatively few standardized instruments have been constructed using these methods and subsequently made available for general use. As a result, looking for a standardized, reliable, and valid instrument to measure a sociological concept is usually frustrating.

To investigate the extent of this problem, all the empirical studies published in two major U.S. sociological journals (American Sociological Review and American Journal of Sociology) and in a methodological journal (Sociological Methods and Re-
MEASUREMENT INSTRUMENTS

REASONS FOR UNDERDEVELOPMENT OF SOCIOLOGICAL MEASURES

The limited production of standard and validated measures of social characteristics may be attributed to a number of causes. One of these is probably a lack of time and other resources for instrument development, validation studies, and replications to cross-validate. Other likely impediments are in the reward structures, opportunities, and constraints of the discipline. These can be illustrated by a comparison with psychology. First, in psychology, there is an institutionalized reward structure for measure development. Journals devoted to psychological measures and the lack of such journals in sociology were noted above. Moreover, there is a large market for psychological tests, and several firms specialize in publishing such tests. It is a multimillion-dollar industry, and authors of tests can earn substantial royalties. Thus, sociology lacks the symbolic and economic reward system that underlies the institutionalization of test development as a major specialization in psychology.

A second reason for the underdevelopment of sociological measures is a situational constraint inherent in much sociological research, especially survey research. Sociologists often must squeeze the measurement of a dozen or so variables into a single hour or half-hour interview. Practitioners, on the other hand, often can use longer, and therefore more reliable, instruments because their clients have a greater stake in providing adequate data and will tolerate undergoing two or more hours of testing. The constraint on the number of survey questions that can be devoted to measuring each variable also may help explain why so few sociologists report reliability coefficients: when the scale is based on only two to four items, the coefficients are likely to be low.

THE FUTURE OF SOCIOLOGICAL MEASURES

There are grounds for optimism and grounds for concern about the future of sociological measurement instruments. The grounds for concern are, first, that most sociological concepts are measured by a single interview question. Second, even when a multiple-indicator instrument is used, it is rarely on the basis of empirical evidence of reliability and validity. Third, both Bonjean, Hill, and McLehose (1967) and Miller (1991) found that the typical measure developed by a sociologist is never used in another study. One can speculate that this hiatus in the cumulative nature of research occurs because of the lack of evidence of reliability or validity, and because authors rarely provide sufficient information to facilitate use of the instrument by others.

The grounds for optimism are to be found in the sizable and slowly growing number of standardized instruments, such as those to measure socioeconomic status (Nam and Powers 1983), juvenile delinquency (Nye and Short 1957), small group interaction (Bales 1950), personality (Borgatta 1964), family interaction (Straus 1990), and alienation (Seeman 1959). The growth of applied and clinical sociology is also a ground for optimism, but one with a certain irony because basic researchers usually believe that they are the guardians of quality in science. In respect to measurement, however, sociological practitioners tend to demand higher-quality instruments than do basic researchers because the consequences of using an inadequate measure are more serious. When a basic researcher uses an instrument with low reliability or validity, it can lead to a Type II error, that is, failing to accept a true hypothesis. This may result in theoretical confusion or a paper not being published (Spanier 1976); but when a practitioner uses an invalid or unreliable instrument, the worst case scenario can involve injury to a client organization or person. Consequently, practitioners tend to demand more evidence of reliability and validity than do basic researchers. As a result, sociologists in applied fields such as family therapy and evaluation research tend to produce and make available a wider choice of more adequate measures. Finally, it can be hoped that an increasing number of sociologists will heed the advice of Blalock (1979, 1982), who has emphasized that inconsistent findings and failure to find empirical support for sound theories in
sociological research may be due to lack of reliable and valid means of operationalizing key concepts in the theories being tested.

(SEE ALSO: Measurement; Reliability; Validity)

REFERENCES


