BECALMED IN MURKY WATERS: MEASUREMENT PROBLEMS IN SOCIOLOGY
Donal E. Muir, University of Alabama

INTRODUCTION. Basic measurement problems continue to plague sociologists. This condition came largely by historical accident, when social reformers captured the infant field. These professionals presented themselves as scientists, but restricted their research efforts almost exclusively to discriptive surveys, relying on nominal and ordinal level measurement of culturally defined and reified social categories. Though such measurement produced data and permitted tests of hypotheses, findings tended to be historical rather than theoretical. This did little to strengthen measurement, and eventually sociologists succumbed to "classical test theory," a set of measurement definitions and notions which, while invalid, enabled researchers to mislead all concerned about the validity of their findings. I will recommend three actions: 1) to adopt the general science measurement model; 2) encourage a reductionist general systems approach to sociological concepts; 3) encourage selective recruitment of sociologists with stronger mathematical and scientific backgrounds.

. Comte, a great admirer of Newtonian mechanics, coined the term sociology "to designate social physics (Comte 1830 42)." But in the United States, sociology was dominated by social reformers who adopted a scientific "front" only after World War I as a strategem for achieving academic recognition (Dynes 1974 173). Many measurement problems were consequently either ignored or so transmogrified that modern researchers may not realize that the original measurement problem remains unsolved.

MEASUREMENT   Measurement is merely a means of storing information about empirical attributes, just as a child's height can be "mapped" to the wall of a house by placing a pencil mark where his head reaches. Such measures can be retrieved later, as when the child is shown how much he has grown. When the "measurement problem" has been solved, there is 1) a measurement operation M which can map any given empirical attribute, e_j to another empirical attribute, n_j for storage, and 2) an inverse operation, M^{-1} which can map the stored representation or measurement back to the original attribute without loss of information (McGinnis 1965 276; Muir 1972 47). Such operations can be represented mathematically as functions which establish a one-to-one correspondence between empirical attributes and measurements:

$$n_j = M(e_j); \qquad e_j = M^{-1}(n_j)$$

LEVELS OF MEASUREMENT
. Most sociologists recognize S.S. Stevens' typology for levels of measurement (1946). 1) If mapping is merely to categories and back, the scale level is nominal; 2) if ranked categories, the scale is ordinal; 3) if to evenly spaced ranked categories, the scale is interval; 4) if to evenly spaced ranked categories with a non-arbitrary zero category, the scale is ratio. The test of scale level is to find what transformations of the scale points can be made, because the category names of nominal, ordinal, interval, and ratio scales are replaceable with those produced by any one-to-one, monotonic, linear, or proportionality transformation, in that order (Coleman 1964 65; McGinnis 1965 278-287).
. Such tests assume that the instrument being evaluated is measuring an empirical attribute that is explicitly theoretically related to the same or more usually a different empirical attribute measured by a second instrument whose scale type has already been established. The evaluation of the scale level of an instru-

ment rests on whether the given transformations interfere with the theory's ability to account for the measurements made by the established instrument. This approach clearly assumes the existence of an instrument with a known scale level, preferably ratio.

. The classic analysis of logic and operations required to establish such primary instruments was made by Campbell (1928). He illustrated with a two-pan balance, and noted how the user could ascertain which of an arbitrary set of masses, such as pebbles, were heavier or equal in mass to others in the comparison operation and then could calibrate the scale by defining the unit of mass as equal to one of a set of equal stones, using two to define two units in the combination operation. Since an absence of pebbles would non-arbitrarily represent zero mass, the user of the scale could conclude that the balance and calibrated stones constituted an instrument of ratio level. Such primary instruments could then be used to evaluate the scale level of other instruments measuring theoretically related empirical attributes. Since Newtonian mechanics relates all of its theoretical concepts such as velocity, acceleration, force, and density to mass, length, & time, primary standard instruments are required only for mass, length, and time.

DISCUSSION.   Campbell's exercise with the balance appears simple enough. Why are not comparable instruments commonplace in sociological research? The early capture of sociology by reformers provides a reasonable thesis. Nominal or ordinal typologies like rich-poor, male-female, and black-white, provided an easy vehicle to communicate the need for reform to the public, and the resulting numbers and associated statistical techniques served to reinforce an image of sociology as a "science actively engaged in research." The era since World War

I can be summarized as the period in which sociology survived as a profession largely by conducting descriptive surveys and gathering case histories of interest to politicians, merchants, and the public, but lacking in scientific interest. This thesis could also account for the continuing emphasis on age, sex, ethnicity, education, occupation, income, and political and religious affiliation even when these variables have no direct theoretical connection with the dependent variables under study (Blalock 1977 6). Given the inertia of social systems, such reports were perhaps useful for short term planning. But they did little to encourage the development of a robust measurement methodology.

. In practice, the opposite occurred. Sociologists increasingly endorsed the invalid notions of "classical test theory" which had originated in psychology (Muir 1977a). This approach to measurement produced a clear break with the established measurement practices of the natural sciences. It also produced such a jumble of basic measurement concepts that today it is often very difficult to exchange even simple measurement ideas among sociologists.

. The frailty of sociological measurement has fostered many individual careers. LaPiere (1969) stated the case: "A concerted re-examination of the whole question of the validity of the research methods that have brought sociology to its present affluence might well bring discredit to a large part of the sociological findings that three decades of prodigious labor have produced, and in the process, undermine the confidence of the general public, the various foundations, and the Federal Government in the fruitfulness of sociology and indeed, of all the social sciences."

. Perhaps there never was much

confidence in sociology. Its non-threatening reputation as a "soft science" made it ideal for sponsored surveys which could be used at the moment to pacify interest groups that "something is being done", to bolster prior decisions, or cheaply gather opinion or demographic data. It is notable that problem-oriented surveys continue as a principal source of funding and opportunity for publication in our contemporary "non-consensual science", giving little impetus for care and concern for measurement.

FIRST RECOMMENDATION
. To strengthen sociological measurement, we must adopt a valid measurement model. Such a model has long been available in the form of the general science measurement model. Suppose we want to evaluate instrument M which on its ith application to a value, e_j of an empirical variable, produces the measurement n_{ij}:

$$n_{ij} = M(e_j)_i$$

To evaluate, we compare the performance of this instrument with that of a standard instrument which, by consensus, defines the parameter value of the empirical variable: $P_j = M_s(e_j)$.

The general science model merely states that any measurement produced by the instrument being evaluated will equal the parameter value plus an error:

$$n_{ij} = P_j + error_{ij}$$

. As simple as it appears, this model is a firm basis for the rigorous definition of fundamental measurement concepts. These definitions relate to point-by-point evaluations of instruments (for over-scale indices, see Muir 1979). The term "E" represents the "expected value", e.g., the arithmetic mean. All necessary calibration procedures are completed before making measurements.

Precision (reliability) = the degree to which the variability of

the theoretical universe of measurements approaches zero. If an instrument is perfectly precise (reliable), each measurement is equal to the expected value of the theoretical distribution: $n_{ij} = E(n_{ij})$

Validity = the degree to which the expected value of the theoretical universe of measurements approaches the parameter value. If an instrument is perfectly valid, the expected value of measurements is equal to the parameter value: $E(n_{ij}) = P_j$

Accuracy = the degree to which the variability and expected value of the theoretical universe of measurements approach zero and the parameter value respectively. If an instrument is perfectly accurate, each measurement is equal to the parameter value: $n_{ij} = P_j$.

. Since these definitions and the basic general science measurement model require standard instruments, it follows that a strengthening of sociological measurement forces us to develop such instruments. This task, from which the puffery of "classical test theory" has too long diverted us, must have highest priority. This undertaking will not be easy or popular, given the present confusion concerning measurement in sociology. Even these three concepts of precision, validity, and accuracy are widely confounded (Muir 1977b).

SECOND RECOMMENDATION
. Because our measurement difficulties go with a lack of effective theory, we should discard the mysticism concerning social systems. More powerful predictive analogs of social systems can be created by viewing them as information processors, rather than as unique entities which possess "free will". Increasing scientific convergence appears in the works of researchers in diverse fields (von Neumann 1958; Homans 1964; Wiener 1965; Skinner 1971; Wilson 1975 and Miller 1978). After we

develop a rigorous theoretical foundation to describe the components for social systems, from ants to humans, we must attempt to model their interactions. Sociologists can profit from general systems theory and specific theories on information processors, such as information theory.

THIRD RECOMMENDATION

. We could assure the more general participation of sociologists in these major scientific developments by recognizing that our measurement theory difficulties stem from measurement complexity, and not from the sui generis character of our subject. Why not selectively recruit future sociologists for their mathematical ability, scientific interest, and intellectual curiosity? This would require reshuffling of professional priorities. It is encouraging to note that the American Sociological Association has taken tentative steps toward accreditation and certification (Footnotes April 1979 1).

REFERENCES

.Blalock H M 1977 Can sociology ever become a discipline? Need for better measurement conceptualization & attention to generalizability. Southern Sociological Society Meeting
.Campbell N R 1928 Measurement & Calculation. London Longmans Green
.Coleman J S 1964 Introduction to Mathematical Sociology. New York Free Press
.Comte Auguste 1830 Cours de Philosophie Positive. In Making of Society, R Bierstedt ed. New York Modern Library
.Dynes, Russell R 1974 Sociology as a religious movement: thoughts on its institutionalization in the United States. American Sociologist 9 169-176 Nov
.Homans G C 1964 Bringing men back in. American Sociological Review 29 809-818 Dec
.LaPiere R T 1969 Comment on Deutscher's looking backward.
American Sociologist 4 41-42 Feb
.McGinnis R 1965 Mathematical Foundations for Social Analysis. New York Bobbs Merrill
.Miller James Grier 1978 Living Systems. New York McGraw Hill
.Muir Donal E 1972 Mathematical systems approach to a synthesis of methodology. Sociological Methods & Research 1 39-63 Aug
_____ 1977a Critique of classical test theory. Psychological Reports 40 383-386 Apr
_____ 1977b Disentangling instrument reliability, validity, and accuracy in the social and behavioral sciences. Sociological Symposium 20 61-72 Fall
_____ 1979 Toward an interscience reconciliation of correlational indices of over-scale precision, validity & accuracy. Southern Sociological Society Meeting
.Skinner B F 1971 Beyond Freedom & Dignity. New York Knopf
.Stevens S S 1946 On the theory of scales of measurement. Science 103 677-680
.von Neumann J 1963 Computer and the Brain. New Haven Yale U Press
.Wiener Norbert 1965 Cybernetics. Cambridge Mass. MIT Press
.Wilson E O 1975 Sociobiology: New Synthesis. Cambridge Harvard U Press

ERRATUM:

In Volume 7, No 2, November 1979, in the Table of Contents the author listing for "Sociology & Social Work: Marriage or Divorce?" should be:
Richard Enos.