Causal Inferences In Behavioral Research

Mildred Field, R.N., Ph.D.
Director of the Graduate Program
School of Nursing
University of Texas Health Science Center at Houston
Houston, Texas

THE ULTIMATE development of nursing models will involve the establishment of causal relationships. The concept of causality underlies much of the conceptual and theoretical development occurring in nursing. Causality thus becomes a concern in nursing research. Postulated causal relations must be tested in empirical research. Twenty of 51 studies which appeared in *Nursing Research*, Volume 27 (1978), for example, attempt to demonstrate a causal relationship. Studies were identified as causal if an explicit statement of causal relationship was made, if the terms *independent variable* and *dependent variable* were used or if a regression analysis was employed.

Within the last decade, there has been increased interest and discussion among behavioral scientists regarding causal analysis. Much of this interest may be attributed to Hubert M. Blalock, Jr. who, in a series of papers, has suggested a process for "making causal inferences from correlation data." Interest in such an approach

0161-9268/79/0021-0081$2.00

raises fundamental questions regarding the concept of causality and the logic which underlies causal analysis.

In delineating the rationale underlying causal inference processes in behavioral research the various methodological strategies can be divided into two categories which, for want of better terminology, are called the traditional and the mathematical model approaches. Such a division is obviously a highly "oversimplified model of reality" and is, in fact, rather arbitrary. Either approach could be subsumed under one or the other category. However, the division seems to reflect a dichotomy in current research concerning the purposes for which data analysis is undertaken and the causal conceptualization which underlies the analysis. Mathematical models have only recently been used in the behavioral sciences and are receiving widespread attention in the literature.

THE CONCEPT OF CAUSALITY

Nagel identifies two major areas of concern in the discussion of causality:

1. the explication of the notion of cause; and
2. the delineation of accepted canons for valid causal inferences.[1]

While the two areas are closely related, the focus here will be on the second area. Criteria for judging that an observed relationship is a causal one must arise from one's conception of cause.

As is frequently pointed out, a variety of causal conceptualizations exists. They reflect individuals' attempts to resolve the series of philosophical issues surrounding the concept of cause. A number of these issues arise from or in response to Hume's classic treatise on causality. Hume maintained that associated with the idea of cause was the observation that one object or event was followed, contiguous in space and time, by another object or event.[2] He noted that although one may observe that B follows A, there is no necessary connection between the two. Instead, imputed connection is the product of psychological habits of thought. Thus the fundamental quality of a causal relation is one only of constant conjunction.

The philosophical issues associated with causality are unresolved and continue to be debated by modern philosophers of science. No definitive resolutions can be provided for each problem. However, fundamental issues associated with the concept of causality include the following:

1. The epistemological or ontological status of causal laws or whether there exist in the real world processes which have correspondence to the cognitive processes.
2. The question of the adequacy of the constant conjunction principle as a description of causal processes. Bunge for example suggests the principle expresses only a relation and ignores the idea that the effect is "produced" by the cause.[3]
3. The question of single or multiple causation. Hume advanced the principle: "The same cause always produces the same effect and the same effect never arises but from the same cause."[2(p170)] Many critics have objected to the artificiality of the princi-

ple and its inadequacy in dealing with empirical phenomena.

4. The question of the existence of a global principle of causality which governs all processes of nature or the existence of specific causal laws. A number of authors seem to agree that causal laws are only one type of lawful regularity in the universe.[3-5]
5. The question of the determinate or indeterminate nature of causal laws. Since the advent of quantum physics, the question has arisen as to whether a state at a particular time completely determines a later state or whether causal laws can only establish dependencies between statistical properties of events.

Despite the unresolved issues, a classical conception of cause in the Humean tradition may be identified. Morgenbesser suggests that the most legitimate use of the term *causal law* is associated with a statement of the type if A then B when:

1. A denotes a type of event that comes just before an event of type B;
2. A and B are events or episodes in bodies or agents that are spatially contiguous and;
3. the occurrence of A is considered a sufficient condition for the occurrence of B even though an event of type B might occur without one of type A preceding it.[6(p120)]

Nagel identifies similar criteria for a causal law but introduces the additional condition of asymmetry.[1]

The response of scientists to these complex and unresolved philosophical issues reflects two major forms of resolution. First, it is suggested by some scientists that the philosophic problems associated with the concept of cause make such analyses tenuous at best and therefore causal inference should be avoided. Dubin, for example, suggests that the term *causal* should not be used.[7] He introduces the concept of a sequential law of interaction which incorporates asymmetric (forcing) and time-ordered relationships. This not uncommon approach would simply seem to relabel the concept of cause. Gibbs, with more direct logic, suggests that current methodologies do not provide an opportunity to test causal propositions.[8] Behavioral research should therefore be directed to the identification of order rather than the demonstration of cause.

It is suggested by some scientists that the philosophic problems associated with the concept of cause make such analyses tenuous at best and therefore causal inference should be avoided.

Advocates of the second form of resolution recognize the philosophical problems associated with the concept of cause but argue that causal imagery implicitly invades scientific thought. Causal analysis must therefore be used as a framework for at least some scientific study. The concept of cause identified and utilized by these authors varies widely, however. Stinchcombe advances a conceptualization which is closely related to the classical concept and incorporates the ideas of association, asymmetry and sufficiency.[9] Rhoads, on the other hand, proposes a concept similar to Bunge's in which the idea of producing and the ontological status, conditional nature and uniformity of causal laws are emphasized.[10]

The conceptualizations of Simon and Blalock provide a major framework for the current interest in causal analysis and undergird the methodological approach described here as model building.[11-13] Simon focuses on the asymmetric quality which he believes characterizes causal laws.[11,12] Causal relations are considered the property of a theoretical model and not a property of the real world. Thus, they have meaning only in terms of the context provided by the theory. Simon identifies the theoretical model as a set of linear equations. The model consists of subsets of equations. Those that can be solved independently of other subsets are termed *self-contained* subsets. A hierarchical ordering of the subsets in the model is assumed. Variables which can be evaluated in one self-contained subset are endogenous to that set and are inserted as exogenous variables in a subset of higher order. A causal ordering of variables is then defined as:

> Let B designate a set of variables endogenous to a complete subset B, and let Y designate the set endogenous to a complete subset C. Then the variables of Y are directly causally dependent on the variables of B ($B \rightarrow Y$) if at least one member of B appears as an exogenous variable in C.[11(p18)]

Operational meaning is given to the system through the fact that each self-contained subset is associated with a specific mechanism of intervention. Thus the nonzero coefficients in the equations can be altered one at a time. The rationale is to limit the flexibility of the mathematical model so it is consistent with an experimental model.[13] The "experimenter" is thus in a position to alter the value of variables which appear at or below the point in which the system is entered.

Blalock has elaborated on Simon's basic work. Simplified models of reality, expressed as a mathematical set of structural equations, are thought to be a bridge between theoretical and empirical languages. Causal relations are specified to be a property of the mathematical model although Blalock maintains some metaphysical assumptions about the correspondence of properties of the model to causal laws in the real world. The recursive form of the equations reflects the inherent asymmetric nature of causal laws. Causal laws describe the process of producing as opposed to mere constant conjunction. (There appears to be some confusion in Blalock's argument at this point. It is stated that causal relations belong on the theoretical level while producing refers to an ontological or empirical process. Producing, however, is included as an essential component of the concept of cause, seemingly mixing the two levels.)

Blalock's and Simon's conceptualizations are unique because the idea of temporal sequence is not included in the exposition of causal relations. Blalock notes that temporal sequence does not encompass the idea of producing and thus "conceptions of cause should not depend on temporal sequence except for the impossibility of an effect preceding its cause."[13(p10)]

In summary, the Simon-Blalock formulation emphasizes an asymmetric conceptualization of cause and effect in which causal relations are defined as properties of the mathematical model. The causal ordering of the variables within the model may be determined by logical or mathematical operations on the model.

THE LOGIC OF HYPOTHESIS TESTING

Causal inference processes can best be understood within the framework of classical scientific inquiry in which hypotheses are proposed and rejected on the basis of empirical evidence. The logic of this process has been fully delineated by Popper who clearly indicates the necessity for a deductive method of hypothesis testing.[14] The problem of induction, well recognized by Hume, invalidates the process of inductive inference. Thus the observation that events of type A have always been followed by events of type B does not justify the conclusion that this sequence of events will continue to be invariant in the future. Neither hypotheses nor theories can be inferred from singular statements.

As Hempel indicates, the deductive method of hypothesis testing rests on the form of logic known as *modus tollens*.[15] The hypothesis (p) is stated and a test implication of the hypothesis (q) is proposed. Thus: If p then q. This form of logic allows for the rejection of the hypothesis. However, a hypothesis can never be proved because the reasoning behind such proof must take the form: If q then p. Such a statement, known as the fallacy of affirming the consequent, is deductively invalid. The hypothesis may or may not be true.

Hypothesis-testing procedures may be understood as attempts to reach a decision regarding the falsity of a singular statement deductively inferred from a universal statement. The falsification of the observation falsifies the theory from which it was logically deduced.

METHODOLOGICAL APPROACHES TO CAUSAL INFERENCES

Two general methodological approaches (the traditional and the mathematical model approaches) related to the study of causal relations among variables are identified. The traditional approach, associated with a classical conception of causality, has as its primary focus the testing of specific hypotheses which postulate a causal relation between variables. The mathematical model approach, though associated with a broader conception of causality by each of its users, incorporates a logical definition of causal ordering within a specified model. While the approach may be used to reject specific causal models, its primary focus is the analysis of the properties and relations within a postulated causal network or system.

The Traditional Approach

EXPERIMENTAL DESIGNS

The paradigm for the testing of causal hypotheses is the exposure of a hypothesis to a situation so designed to allow for its disconfirmation. The classical design for such purposes is the experimental design proposed by Fisher.[16]

Leonard compares the rationale of the experimental method to the deductive method of hypothesis testing.[17] He notes that to infer that one variable, X, has a causal effect on a second variable Y, three assumptions must be met:

1. X and Y vary together;
2. X precedes (or occurs simultaneously with) Y and;
3. all other antecedent independent variables are taken into account.

These assumptions may be symbolically represented and inserted into the premise of the argument:

1. $X \rightarrow Y$
2. $X < Y$ $\Big\rangle\, p$
3. AOAIVTIA

The test implication of the hypothesis may be represented as $(rXY \neq 0)$. The assumption of simple (singular) causation could be represented as $rXY = 1.00$. However, if multiple causation is assumed, the correlation between X and Y cannot be expected to be unity. Empirical observation may be made and if not q $(rXY \neq 0)$ then not p $(X \rightarrow Y,\ X < Y,$ AOAIVTIA).

In testing a hypothesis only one assumption can be tested. In the above case, if not q then any of the three assumptions may be false. It is impossible to conclude that only $(X \rightarrow Y)$ is false. The presence of several untested assumptions allows for the possibility that rival hypotheses may plausibly account for the result that $(rXY \neq 0)$. For example, if assumption 2 is not previously confirmed as true then it is plausible that Y determines X. If the truth of assumption 3 is not predetermined, then the possibility of a spurious correlation between X and Y, caused by a common third variable, is a plausible rival hypothesis.

An experimental design renders implausible rival hypotheses. The better the experimental design, the fewer the plausible rival hypotheses that remain. In an experimental design, the condition of temporal order is achieved by manipulation of the independent variable. The treatment group(s) receive differing amounts of the independent variable than the control group(s) and, following the manipulation, measurement is made of the dependent variable.

An experimental design renders implausible rival hypotheses. The better the experimental design, the fewer the plausible rival hypotheses that remain.

The condition that other variables are taken into account may be achieved in either of two ways. Variables which are believed to affect the relation between X and Y may be controlled so they do not vary during the experiment. It is frequently impossible to either identify or control all such variables. Thus the strength of the experimental design lies in randomization, the second method of accounting for other variables. In the process of randomization, individuals are randomly assigned to either treatment or control groups. Such a procedure ensures that the groups differ initially only by chance, i.e, that there are no systematic differences between the groups which may account for the outcome of the experiment. A probability calculation or test of significance is then utilized to indicate the probability that the observed relationship between the variables was due to randomized variables that were not equalized.[18]

QUASI-EXPERIMENTAL DESIGNS

The practicality of achieving sufficient control to allow for experimental manipulation of the independent variable and randomization of subjects is limited in behavioral research. One approach to this problem is to design situations which closely approximate an experimental design and to use the logic underlying such designs with the explicit realization of the shortcomings of each design.

Campbell and Stanley have been major advocates of this approach. They suggest that since hypotheses can never be proved but only probed, quasi-experimental designs offer sufficient probing ability to be of value when no better designs are available.[19]

Although Campbell and Stanley present ten examples of quasi-experimental designs, they emphasize the construction, by individual investigators, of creative data collection arrangements which capitalize on the unique features of specific research settings.

To this end they advocate the use of both experimental and natural manipulations of the independent variable. They note, however, that designs which incorporate naturally occurring manipulations have less validity and, to some critics, questionable status as an experimental design. The designs presented differ widely. However, they have in common certain features which strengthen their ability to probe causal hypotheses. At a minimum, all the designs occur over time, allowing the researcher to determine the temporal order of the variables. Some of the stronger designs incorporate randomization. When randomization is not possible, the assumption that there is no systematic difference between control and treatment groups is strengthened by the administration of pretests. Control groups or repeated measurements over periods of time serve as a baseline of comparison for treatment effects and render several rival hypotheses implausible. Various combinations of these techniques are used in specific situations to increase the validity of the design.

Evaluation of the designs is made possible through a checklist of major sources of invalidity or plausible rival hypotheses. The majority of the designs presented approach or meet the internal validity of the true experimental design. The external validity or generalizability of the results to other populations or settings is lower due to the inability to control for interaction effects inherent in the experimental situation.

Mathematical Model Approach

Behavioral research in the last ten years has been marked by the introduction of a number of techniques for the analysis of cross-sectional correlation data. These techniques, appropriate for quantitative data, come principally from econometrics and biometrics and include causal inference procedures[13] and path analysis.[20]

These procedures will be considered within the general framework of mathematical models.[21]

EXPLICATION AND ASSUMPTIONS OF MATHEMATICAL MODELS

Land has provided an explication of mathematical models and their relationship to formal theory and empirical data.[22] Basically, a mathematical model may be defined as the specification, in mathematical terms, of a verbal theory. Such specification includes variables and free parameters which correspond to theoretical concepts and equations which correspond to the axioms of the theory. Variables within the model are solved for as functions of the parameters, termed free parameters since they may be fixed and manipulated by the model builder. The model in this form thus provides a general description of the class of events associated with the verbal theory since parameters—and hence variables—may take on values within a broad range. The mathematical model is therefore "an expression of the theory in a form that permits analysis of its logical implications."[22(p184)]

A description of a specific set of events must include observations of the conditions which are obtained in that particular instance. The process of model solution yields a set of equations known as the reduced form of the mathematical model. These equations allow one to estimate the parameters of the model from empirical data and to incorporate in the model conditions specific to empirical events. The reduced form equations do not represent specific causal mechanisms and are not necessarily independent of one another. Thus it is necessary to utilize structural-parameter estimation procedures to obtain the mathematical structure of the model. The mathematical structure of the model incorporates both the structural equations of the mathematical model and the estimated values of the structural parameters from the empirical data and, in this form, allows for the delineation of the specific causal relations between variables within the model.

It may further be noted that mathematical models may be classified as one of four types.[22] Static models have no reference to time while dynamic models relate to values of parameters and variables which are measured at different points in time. Exact models describe the relations between variables without error while stochastic models incorporate disturbance terms which reflect the presence of factors acting randomly on the model.

Behavioral research is characterized by the use of stochastic models. This reflects the fact that causal laws apply exactly only under ideal conditions. As Blalock points out, sociologists must deal with real life situations and cannot evaluate causal relations under ideal conditions.[13,23] Stochastic models which allow for random disturbance in the system may therefore be used to evaluate causal relations in empirical situations. The majority of models used in sociological research have been static, although some recent progress has been made in the development of dynamic models.

Generically, mathematical models may be identified as structural equation models. Because the recursive model is believed adequate for explicating the logic underlying causal inference processes, it will be used here as a point of reference.

As Heise notes, a valid mathematical model provides a means of explanation and prediction.[21] Land argues that since social scientists are unable to experiment,

Land argues that since social scientists are unable to experiment, mathematical models provide ways to test theories in which conditions of observation can be incorporated.

mathematical models provide ways to test theories in which conditions of observation can be incorporated.[22] The point must be emphasized that procedures do not provide a magical method of "deducing causal relations from the value of correlation coefficients."[24] In fact, the procedures rest on a number of untestable a priori assumptions regarding the causal relations presumed to hold in the model.

Heise has clearly specified the nature of these assumptions for recursive models.[21] The data used in such analyses must meet the usual requirements for regressive analysis and must be based on measurements that have high reliability and validity. For the recursive model, the theory must specify relations which are linear and additive and which involve only one-way causation. Of particular importance in the causal interpretation of such data are the assumptions regarding the disturbance terms and the causal ordering of the variables.

In recursive models, the usual assumption is that the disturbance terms of the dependent variables are uncorrelated with each other and with system inputs, where inputs are defined as variables directly or indirectly affecting two or more dependent variables. The disturbance terms in the model reflect variance which is unaccounted for by variables in the system and may be conceptualized as the effect of unmeasured variables on the system. As Heise points out, a correlation between disturbance terms indicates the presence of unmeasured variables affecting two or more dependent variables. The assumption that the disturbance terms are not correlated is thus equivalent to the assumption that all system inputs are included in the model.

To construct a mathematical model, it is necessary to specify which variables are the dependent and which are the input variables and to further specify the correct ordering of the dependent variables. These two assumptions thus require extensive a priori knowledge of the problem from theoretical or empirical work. Insufficient a priori knowledge leading to specification errors in the model will, as Bohrnstedt and Carter demonstrate, result in meaningless parameter estimations.[25]

USES OF MATHEMATICAL MODELS IN BEHAVIORAL RESEARCH

The stringency of such requirements has led many to question the applicability of such techniques in behavioral research. Their use is obviously optimum in situations associated with well developed causal theory. However, these rarely exist in behavioral research. The major application of such techniques in social research has been by Otis Dudley Duncan who, in a series of papers, has developed and explored social stratification models. These models take advantage of the temporal order of events in the life cycle of individuals and thus incorporate a strong a priori basis of causal ordering.

Through these papers, Duncan has demonstrated many of the advantages and uses of mathematical models. Primary among these are the explicit delineation of assumptions and the internal consistency required by mathematical models. As Duncan notes in his initial paper on path analysis as a pattern of interpretation, path analysis "is invaluable in making explicit the rationale for a set of regression calculations."[20(p7)]

Duncan's basic work on stratification involved the delineation of a basic model of stratification and parameter estimation procedures for a sample of 20,000 American men.[26] He provided an interpretation of the stratification process in American society as well as numerical estimates of the structural parameters applicable to a large segment of the population.

Duncan has extended this basic work and, in the process, has demonstrated the further value of mathematical models in situations where the causal ordering of the variables may involve ambiguity. The basic purpose of such analyses is, as Wright comments, "to find the logical consequences" of any particular set of causal assumptions.[24] In the first example, the problem of unmeasured theoretical variables is considered.[27] A speculative analysis is undertaken which provides information regarding areas of difficulty in analysis and interpretation. Such a technique can be useful in identifying and delineating problems for further research. The second example, extending much beyond the simple recursive model, incorporates reciprocal causation and further addresses the problem of measurement error.[28] With no firm basis for the construction of models which incorporate social-psychological variables, it is suggested that one construct reasonable models and consider their plausibility in terms of the theory and data. Such an approach makes explicit the assumptions of the model and presents "interpretations in such a form that their weaknesses and those of the theories which give rise to them are fairly evident."[28(p137)]

The initial literature on mathematical models emphasized model-testing procedures as an approach to making causal inferences from correlation data. Blalock has been a principal advocate of this approach. He argues that the logic underlying causal inference procedures from experimental and nonexperimental data differs only in degree and that techniques associated with mathematical models may be used to test the adequacy of proposed causal models.[13] While the logic may be the same, the degree of difference would seem to be great. The primary difference relates to the stringency of test conditions for each type of data.

In experimental situations, the effect of confounding variables is taken into account through the design of the experiment and the randomization process. In causal model-testing procedures, control of confounding variables is an assumption incorporated in the model, i.e., that all input variables have been included in the model. Blalock notes that randomization eliminates the effects of only those variables which are properties of the system at the time of observation (property variables) but cannot eliminate the effects of variables which impinge on the system from the outside environment (forcing variables). He therefore argues that in both experimental and nonexperimental research, a point is reached where one has to make an assumption regarding the effects of confounding variables. A well designed experiment, however, includes the identification and control of such forcing variables. Campbell and Stanley's checklist of sources of invalidity, for example, includes such general factors as history and testing effects.[19] Thus, while it is true that assumptions must inevitably be made, such assumptions would seem better supported in experimental than nonexperimental research.

Blalock's model-testing procedure is applicable only in the case of overidentified models. A full scale recursive model will provide no basis for the rejection of the model. The empirical correlations will always be reproduced. (Overidentification refers to a situation where there is more than enough information from the data to estimate the structural parameters.[29] In the case of a recursive path model, this would amount to the specification that one or more possible paths between variables is absent.) The excess equations which result from the overidentification of the model may be used as prediction equations. In Blalock's formulation, this amounts to the prediction that particular partial correlations will be zero. (Blalock's causal inference procedure is a weak form of path analysis and Heise[21] has reinterpreted the causal inference model-testing procedure in terms of path analysis.) If the sample values are close to zero, the model is retained. If the sample values differ from zero and the discrepancy is too large to be attributed to sampling error, the model is rejected.

Overidentification refers to a situation where there is more than enough information from the data to estimate the structural parameters.

The procedure does not, however, help in the problem of specification of the causal ordering of the variables. Acceptance of the model does not provide evidence that the causal ordering in the model is the correct one since any number of alternative models may lead to the same predictions. Such an assumption would correspond to the fallacy of affirming the consequent.

Heise has been most critical of the Blalock procedure.[21] He notes that in using the method to compare the adequacy of two theoretical models, one may appear more adequate but neither may be right. He further points out that the closer the model is to a complete recursive system, the better chance it has to be judged adequate, though it may be further from the "true" model.

An alternative and more circumscribed approach to model testing involves testing specific portions of a postulated model. The question of interest involves the presence or absence of linkages between certain variables in the model. Goldberger has described such an approach as the testing of overidentified restructions of the model.[29] The overidentified restructions serve as null hypotheses which can be probed by significance tests. Duncan describes one such test and indicates that considerations regarding revision of the model rest on both statistical and substantive information.[20] Goldberger notes that such testing procedures may be used in a constructive way to revise portions of the model rather than as a basis for the rejection of the entire model[29] and Heise comments on the usefulness of such an approach in developing empirically based theory.[21]

EXPERIMENTATION AND MODEL BUILDING

The experimental design remains the ideal approach to testing causal hypotheses. The limitation of the design, other than its feasibility, is its restriction to situations involving few variables. Mathematical models, on the other hand, may be used to analyze causal relations in a multivariate system, although their use rests on a number of untestable assumptions. A combination of the two approaches may therefore provide a reasonable strategy for sociological research.

As Heise suggests, a series of experimental designs may be used to construct mathematical models.[21] Ideally, these would be used to determine both the form and parameters of the model but, at a minimum, they could be utilized as a basis for the a priori assumptions regarding causal ordering of the variables. It is questionable whether this approach would solve the problems of causal inference in behavioral research, for it would seem that the causal relations between two variables might be altered when introduced into a multivariate system. In view of the current alternatives, however, a carefully conceived program of experimentation and model building might provide a more rigorous approach to the analysis of causal relations.

REFERENCES

1. Nagel, E. "Types of Causal Explanations" in Lerner, D. ed. *Cause and Effect* (New York: The Free Press 1965) p. 11.
2. Hume, D. *A Treatise on Human Nature, Volume I* (London: Everyman's Library 1911).
3. Bunge, M. *Causality* (Cambridge: Harvard University Press 1959).
4. Nagel, E. *The Structure of Science* (New York: Harcourt, Brace and World 1961).
5. Kaplan, A. "Noncausal Explanation" in Lerner, D. ed. *Cause and Effect* (New York: The Free Press 1965) p. 145-156.
6. Morgenbesser, S. "Scientific Explanation" in *The International Encyclopedia of the Social Sciences* (New York: Macmillan 1968).
7. Dubin, R. *Theory Building* (New York: The Free Press 1969).
8. Gibbs, J. "Causation and Theory Construction." *Social Science Quarterly* (March 1972) p. 815-826.
9. Stinchcombe, A. *Constructing Social Theories* (New York: Harcourt, Brace and World 1968).
10. Rhoads, J. "On Causality in Social Science" *Sociol Inquiry* 41 (Winter 1971) p. 27-37.
11. Simon, H. *Models of Man* (New York: John Wiley & Sons 1957).
12. Simon, H. "Causation" in *The International Encyclopedia of the Social Sciences* (New York: Macmillan 1968).
13. Blalock, H., Jr. *Causal Inferences in Nonexperimental Research* (New York: W. W. Norton and Co. 1964).
14. Popper, K. *The Logic of Scientific Discovery* (London: Hutchinson & Co. 1968).
15. Hempel, C. *Philosophy of Natural Science* (Englewood

Cliffs, N.J.: Prentice-Hall 1966).
16. Fisher, R. *Design of Experiments* (Edinburgh: Oliver and Boyd 1935).
17. Personal communication with R. Leonard, Ph.D., Professor, Department of Sociology, University of Arizona. 1971.
18. Wooldridge, P., Leonard, R. and Skipper, J. *Methods of Clinical Experimentation to Improve Patient Care* (St. Louis: The C.V. Mosby Co. 1978).
19. Campbell, D. and Stanley, J. *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand McNally and Co. 1966).
20. Duncan, O. D. "Path Analysis: Sociological Examples." *Am J Sociol* 72 (July 1966) p. 1-16.
21. Heise, D. "Problems in Path Analysis and Causal Inference" in Borgatta, E. ed. *Sociological Methodology 1969* (San Francisco: Jossey Bass 1969).
22. Land, K. "Formal Theory" in Costner, H. ed. *Sociological Methodology 1971* (San Francisco: Jossey Bass 1971) p. 175-220.
23. Blalock, H., Jr. "Theory Building and Causal Inferences" in Blalock, H. and Blalock, A. eds. *Methodology in Social Research* (New York: McGraw-Hill Book Co. 1968) p. 155-198.
24. Wright, S. "The Method of Path Coefficients." *Ann Math Stat* 5 (September 1934) p. 161-215.
25. Bohrnstedt, G. and Carter, T. "Robustness in Regression Analysis" in Costner, H. ed. *Sociological Methodology 1971* (San Francisco: Jossey Bass 1971) p. 118-146.
26. Blau, P. and Duncan, O. *The American Occupational Structure* (New York: John Wiley & Sons 1967).
27. Duncan, O. "Contingencies in Constructing Causal Models" in Borgatta, E. ed. *Sociological Methodology 1969* (San Francisco: Jossey Bass 1969) p. 74-112.
28. Duncan, O., Archibald, O. and Porter, A. "Peer Influence on Aspirations: A Reinterpretation." *Am J Sociol* 74 (1968) p. 119-137.
29. Goldberger, A. "Structural Equation Models: An Overview" in Goldberger, A. and Duncan, O. eds. *Structural Equation Models in the Social Sciences.* (New York: Seminar Press 1978) p. 1-18.