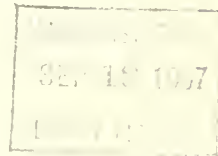


TOWARD A NON-EXPERIMENTAL METHOD FOR CAUSAL
ANALYSES OF SOCIAL PHENOMENA

George F. Farris

August, 1967



#67-279

W 208
.701414
OCT 29 1967

RECEIVED
OCT 2 1967
M. I. T. LIBRARIES

Abstract

A method is described for detecting causal relationships among social phenomena in the natural setting. Based primarily upon obtaining an association between two factors, one measured earlier in time than the other, and controlling the influences of outside factors, it considers causality to be a matter of degree and allows for symmetrical causal relationships. A study of organizational behavior is described to demonstrate the applicability of the method. After contrasting the method to path analysis and panel-study analysis, it is suggested that behavioral scientists think in terms of three levels of research on complex cycles of social behavior: one point-in-time correlations to determine general associations among factors, methods like the one proposed here to determine general causal relationships, and, when possible, experiments to determine more precisely the causal relationships among factors of particular interest.

TOWARD A NON-EXPERIMENTAL METHOD FOR CAUSAL ANALYSES OF SOCIAL PHENOMENA

George F. Farris¹

Massachusetts Institute of Technology

An important goal of the social sciences is to determine causal relationships among phenomena which occur in natural settings. Extensive use has been made of two kinds of studies to achieve this goal: the correlational and the experimental. Correlational studies have been especially useful in describing relationships among factors in the natural setting. Questionnaire surveys, for example, have found significant relationships between a number of organizational factors and performance of people in ongoing organizations. The usual correlational study, however, is not very useful in determining cause and effect.

Experiments have been advocated as a way to detect causal relationships. Although they can be very useful for this purpose, they, too, have their limitations. After a very persuasive evaluation of sixteen common experimental designs, Campbell and Stanley (1963), for example, classified only three as "true" experimental designs. The difficulties in using these "true" experimental designs in the natural setting are formidable. Ethical and power reasons are often prohibitive, and even tightly controlled one-way experiments deal with only a small portion of the factors which may be involved in the actual causal relationship.

¹This paper is based upon the author's dissertation, submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree at the University of Michigan. The author is grateful for the comments and suggestions of the members of his committee: Robert L. Kahn, Chairman; Frank M. Andrews, Basil S. Georgopoulos, Abraham Kaplan, and J. E. Keith Smith. Part of the research was supported by grant NSG-489-28-014 from the National Aeronautics and Space Administration.

Moreover, experiments typically investigate only one side of what may be the actual causal relationship. In predetermining the "independent" and "dependent" variables they fail to investigate the possibility of the dependent variable affecting the independent. Although this possibility is typically excluded by the closed system of the experimental design, it may actually occur in the "real world" situation which the experiment purports to mirror. For example, performance has been studied in experiments as a function of type of supervision. In the natural setting, it is also conceivable that the type of supervision is dependent upon the individual's performance. When correlational findings are interpreted in conjunction with findings from laboratory experiments, the risky assumption is made that the factors studied in the laboratory and the field reflect the same phenomena.

In the present paper an attempt is made to develop a non-experimental method for making causal analyses. If such a method could be developed, it might have the potential of combining the major advantages of the correlational and experimental methods. Like correlational methods, it might be readily usable in natural settings, and, like experiments, it might allow causal explanations to be made. At the heart of the method is a working definition of a causal relationship based primarily upon obtaining an association between two factors, one measured earlier in time than the other, and controlling the influences of outside factors. Relying upon the common-sense notion that an effect cannot precede its cause in time, the working definition considers causality to be a matter of degree and allows for symmetrical causal relationships. In terms of this definition it is possible to define causal, intervening, and resultant factors and to specify possible patterns of causal relationships among a number of factors.

The method proposed here turns out to share some characteristics of non-experimental methods for causal analysis proposed by others, notably path analysis and panel-study analysis. After describing the method and an empirical test of it, we shall contrast it to these other approaches.

DEFINITION OF A CAUSAL RELATIONSHIP

Let us consider two factors, X and Y, measured at the same point in time. For purposes of illustration, let X refer to closeness of supervision and Y refer to production. Let us assume that X and Y are found to be associated. What does this fact tell us about a possible causal relationship between them? By itself the association tells us nothing about causality. It could be due to a number of different causal relationships. I believe that the list below exhausts the possibilities.

Case 1. X may cause² Y. General supervision may cause high production.

Case 2. Y may cause X. High production may cause general supervision.

Case 3. X may cause Y and Y may cause X. In chicken-and-egg like cyclical fashion, close supervision may cause low production and low production in turn may cause close supervision. More specifically, events like the supervisor's looking over the subordinate's shoulder may cause the subordinate to decrease his production when the supervisor is not looking over his shoulder, which in turn causes the supervisor to look over the subordinate's shoulder even more frequently.

Case 4. The association between X and Y may be caused by other factors which cause both X and Y. Inexperience of the subordinate in his job is an

²Throughout this paper I shall use the term "cause" in the sense "is a cause of." To say, "X causes Y" does not imply that some factor other than X may not also be a cause of Y. For example, both "general supervision" and "diversity of technical assignments" may cause production of scientists in the sense in which I am using the term "cause."

example of such a factor. Inexperience may cause a supervisor to supervise more closely and at the same time may cause the subordinate to produce less. In this case closeness of supervision and production would be associated although neither would cause the other. A more vivid illustration occurs in the case of an apparently paradoxical positive association between the frequency of visits of physician and the number of deaths of their patients. (Hopefully) neither factor is causally related to the other. Rather, each factor is caused by an additional factor: the severity of the patient's illness.

Given this number of possible relationships when two factors are associated, how then are we to determine which case or cases hold true? Let me propose a two-step process for doing this:

1. Make sure that the association between X and Y is not due to the influence of a factor other than the two being considered. Accomplishing this step is not difficult to do in theory. Statistical techniques such as partial correlation and multiples regression analyses allow the association between two factors to be determined while controlling for the effects of additional factors. In practice, however, things are not quite so easy. Rarely can we be absolutely certain that we have identified all the potentially contaminating factors to control statistically. After this step has been accomplished, the association between X and Y may remain, be reduced substantially, or even reverse in sign. If the association between X and Y remains substantial after removing the effects of additional factors, it cannot be due to the operation of causal relationships pointed to in Case 4, at least with respect to those factors examined. Only cases 1,2, or 3 can obtain. To distinguish among them, a second step is necessary.

- 2.³ Measure both X and Y at two points in time. Examine the strengths of the association between:

³ Step 2 turns out to be very similar to methods proposed earlier by Lazarsfeld (1946), Campbell and Stanley (1963), and Pelz and Andrews (1964).

(a) X measured at Time 1 and Y measured at Time 2, and

(b) Y measured at Time 1 and X measured at Time 2.

If (a) is substantially different from zero, then Case 1 holds (X causes Y).

If (b) is substantially different from zero, then Case 2 holds (Y causes X).

If (a) and (b) are both substantially different from zero, then Case 3 holds (X causes Y and Y causes X).

Let us examine Step 2 more closely, referring to Figure 1.

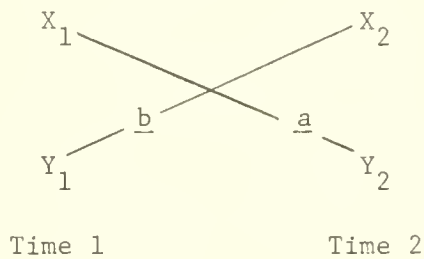


Figure 1. Cross-lagged associations between two factors measured at two points in time.

Step 2 rests upon a simple, common-sense assumption: a cause precedes its effect in time, but an effect cannot precede its cause in time. Thus, closeness of supervision measured at Time 1 can cause production measure at Time 2, but production measured at Time 2 cannot cause closeness of supervision as measured at Time 1. Similarly, production measured at Time 1 can cause closeness of supervision to be what it was at Time 2, but closeness of supervision at Time 2 cannot cause production to be what it was at Time 1. Association a measures the causal influence of X on Y, while association b measures the causal influence of Y on X. If a is substantially greater than zero, then X causes Y (Case 1). If b is substantially greater than zero, then Y causes X (Case 2). If a and b are both substantially greater than zero, then X causes Y and Y causes X. (Case 3).



(a) X measured at Time 1 and Y measured at Time 2, and

(b) Y measured at Time 1 and X measured at Time 2.

If (a) is substantially different from zero, then Case 1 holds (X causes Y).

If (b) is substantially different from zero, then Case 2 holds (Y causes X).

If (a) and (b) are both substantially different from zero, then Case 3 holds (X causes Y and Y causes X).

Let us examine Step 2 more closely, referring to Figure 1.

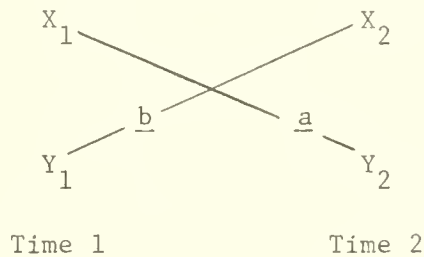


Figure 1. Cross-lagged associations between two factors measured at two points in time.

Step 2 rests upon a simple, common-sense assumption: a cause precedes its effect in time, but an effect cannot precede its cause in time. Thus, closeness of supervision measured at Time 1 can cause production measure at Time 2, but production measured at Time 2 cannot cause closeness of supervision as measured at Time 1. Similarly, production measured at Time 1 can cause closeness of supervision to be what it was at Time 2, but closeness of supervision at Time 2 cannot cause production to be what it was at Time 1. Association a measures the causal influence of X on Y, while association b measures the causal influence of Y on X. If a is substantially greater than zero, then X causes Y (Case 1). If b is substantially greater than zero, then Y causes X (Case 2). If a and b are both substantially greater than zero, then X causes Y and Y causes X. (Case 3).

What may we conclude if X and Y are associated, but neither a nor b shows associations large enough to indicate that X causes Y or Y causes X? The most probable inference is that the interval chosen between Time 1 and Time 2 is not one which captures the dynamics of the causal relationship. That is, the time lag needed for X to affect Y or vice versa is longer or shorter than the interval chosen between Time 1 and Time 2.⁴

Let us digress for a moment at this point to mention a few philosophical aspects of this procedure for inferring causality. One may wonder how an event which occurred in the past may affect something which happens in the present. Our answer to this question is that some "trace" of the past event or cause must still be present in some form at the time the effect occurs. For example, a "trace" of last Tuesday's pat on the back affects Friday's production. The manner in which this occurs is unclear, and we do not wish to speculate about it at this point. Secondly, the fact that X has led to Y in the past may create an expectation that X will lead to Y in the present and future. For example, the fact that high performance has led to a bonus in the past may create the expectation that high performance will lead to a bonus in the future. In a sense, such an expectation may be a cause of performance: because he expects a bonus, a person produces more. Although the two-step approach does not account for expectations specifically, it could conceivably do so if a person's expectations were measured at two points in time, just as factors X and Y were in Figure 1. Thirdly, when we say that general supervision causes high production, we are talking about factors or concepts abstracted from events such as looking over the subordinate's shoulder. Strictly speaking, causal relationships occur among these events and not the factors abstracted

⁴You will recall that throughout our discussion of Step 2 we have been dealing with a situation in which X and Y are associated and the influence of additional factors on this association has been controlled by Step 1.

from them. When we say that general supervision causes high performance, we mean that events underlying the factor general supervision cause events underlying the factor high performance.

Now we are prepared to advance a working⁵ definition of a causal relationship on the basis of Steps 1 and 2:

If two factors, X and Y, are each measured at two different points in time, Time 1 and Time 2; if X measured at Time 1 is associated with Y measured at Time 2 (a in Figure 1) or vice versa (b in Figure 1); and if their association is not due to the influence of factors other than X and Y, a causal relationship exists between X and Y.

We may distinguish two types of causal relationships:

Type I--A Type I causal relationship occurs if X causes Y more than Y causes X or vice versa. A Type I causal relationship occurs if either of two sets of conditions holds:

Type Ia⁶--X causes Y. X causes Y in a Type Ia causal relationship if the association (a in Figure 1) between X measured at Time 1 and Y measured at Time 2 is substantially different from zero, and the association (b in Figure 1) between Y measured at Time 1 and X measured at Time 2 is not substantially different from zero. (Y causes X in a Type Ia causal relationship if association b is substantially different from zero and association a is not.)

Type Ib--X causes Y and Y causes X, but one of the X-Y cross-lagged associations is substantially greater than the other.

Type II⁷--A Type II causal relationship occurs if X causes Y and Y causes X, and the X-Y cross-lagged associations are not substantially different in magnitude.

⁵This definition is not intended to shake any philosophical foundations; rather, it purports to be a definition on which the behavioral scientist can find handles for his research without unduly violating the philosopher's notions of cause.

⁶ What we have called a Type Ia causal relationship here is also known as an *antisymmetric* relationship.

⁷ What we have called a Type II causal relationship here is also known as a *symmetric* relationship.

THE FRAMEWORK

On the basis of our working definition of a causal relationship, we are prepared to consider three kinds of factors: causal, intervening, and resultant. After defining each we shall spell out the possible patterns of associations among three factors measured at three points in time, paying particular attention to those which demonstrate the existence of causal, intervening, or resultant factors.

Causal, intervening, and resultant factors defined

Let us define causal, resultant, and intervening factors in terms of our definition of a causal relationship. If two factors are causally related to each other, the causal factor or cause is the factor which causes the second factor, and the resultant factor or effect is the factor which is caused by the first factor.⁸ If X and Y are related so that X causes Y, then X is the causal factor and Y is the resultant factor.

The definitions of causal and resultant factors mean that X and Y must be measured at different points in time, the influence of third factors must be controlled, and the "across-time" or "cross-lagged" association between X and Y must be different from zero by an amount large enough for us to conclude that it is a substantial one. The factor which was measured earlier in time is then the causal factor, and the one measured later is the resultant factor. For example, in Figure 1, X is a causal factor and Y is a resultant factor if association a is substantially different from zero.

A factor which is a resultant factor in a causal relationship with one factor and a causal factor in a causal relationship with a second factor is

⁸ What we have called a causal factor corresponds to the common sense notion of a "cause," and what we have called a resultant factor corresponds to the common sense notion of an "effect." We shall use the terms interchangeably.

an intervening factor between the first and second factors. If Y is a resultant factor in a causal relationship with X, and Y is a causal factor in a causal relationship with Z, then Y is an intervening factor between X and Z.

Operationally, for Y to be an intervening factor between X and Z, two things must occur. First, X must be measured at a point in time earlier than that at which Y is measured, the influence of third factors must be controlled, and the across-time association between X and Y must be substantially different from zero. Secondly, Y must be measured at a point in time earlier than that at which Z is measured, the influence of third factors must be controlled, and the across-time association between Y and Z must be substantially different from zero. X may or may not cause Z directly.

We should note here that these definitions allow a particular factor to be a causal factor in one instance and an intervening or resultant factor in another. The definitions do not indicate whether Type I or Type II causal relationships are occurring. If X and Y are associated in a symmetrical causal relationship, for example, it is entirely possible for X to be both a causal and a resultant factor with respect to Y. In the next section we shall treat this point more fully.

Patterns of association

The above definitions of causal, intervening, and resultant factors are based upon patterns of association among factors measured at different points in time. Similarly, we distinguish between Type I and Type II causal relationships on the basis of patterns of association among factors measured at different points in time. Let us turn to an examination of such patterns of association. First, we shall specify patterns which indicate causal, intervening, or resultant factors. Then we shall discuss the problem of identifying Type I and Type II causal relationships.

Patterns which identify causal, intervening, and resultant factors.

Figure 2 depicts three factors, X, Y, and Z, and the times, T_0 , T_1 , and T_2 at which they are measured.

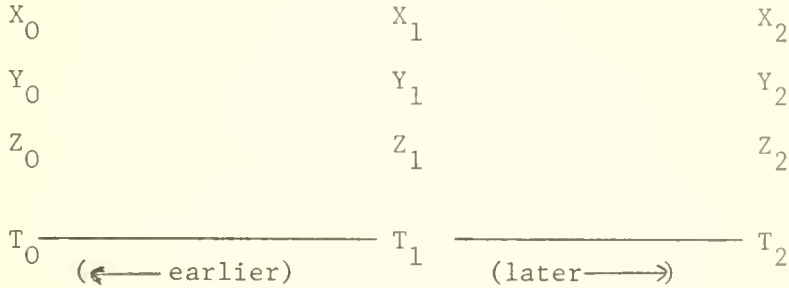


Figure 2. Three factors, each measured at three points in time.

In order to specify the possible associations among the three factors, we shall select a subset of them consisting of each factor measured at one point in time, X_0 , Y_1 , and Z_2 . The possible associations within this subset are representative of the possible associations within any subset of three factors, each of which is measured at a different point in time. The associations between X_0 , Y_1 , and Z_2 are illustrated in Figure 3. The association between X and Z is symbolized by a, the association between X and Y is symbolized by b, and the association between Y and Z is symbolized by c.

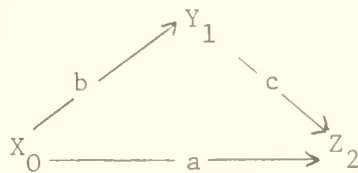


Figure 3. Associations among three factors, each measured at a different point in time.

Each of these associations may be either substantial or insubstantial in size and, if substantial, it may be either positive (a high score on X is

associated with a high score on Y) or negative (a high score on X is associated with a low score on Y) in direction. Let us label a substantial association in either a positive or a negative direction with a plus (+) sign as a superscript (e.g., a^+), and an association which is not substantial with a zero (0) as a superscript (e.g., \underline{a}^0).

Since each association can be either substantial or insubstantial, there are 2^3 or eight patterns of relationships possible between X, Y, and Z. Let us consider each of these in turn. It may be easier to follow this discussion if X, Y, and Z are thought of as referring to specific factors. For example, it may help to consider X as referring to closeness of supervision; Y, to involvement in one's work; and Z, to production. In each of these patterns we shall assume that the relationships are not affected by factors other than X, Y, and Z.

Pattern 1. $\underline{a}^+, \underline{b}^0, \underline{c}^0$. X causes Z. For example, general supervision causes high production but not via involvement. X is a causal factor and Z is a resultant factor.

Pattern 2. $\underline{a}^0, b^+, \underline{c}^0$. X causes Y. For example, general supervision causes involvement but involvement does not cause production, nor does general supervision cause production. X is a causal factor, and Y is a resultant factor.

Pattern 3. $\underline{a}^0, \underline{b}^0, \underline{c}^+$. Y causes Z. For example, involvement causes production but general supervision does not cause involvement. Y is a causal factor, and Z is a resultant factor.

Pattern 4. $\underline{a}^0, \underline{b}^+, \underline{c}^+$. X causes Y and Y causes Z. Y is an intervening factor between X and Z. For example, general supervision causes involvement, and involvement causes high production. General supervision does not influence production directly as in Pattern 1. Rather, general supervision influences

production indirectly through a causal path in which it directly influences a factor, involvement, which is a cause of production. X is a causal factor; Y is a causal, resultant, and intervening factor, and Z is a resultant factor.

Pattern 5. \underline{a}^+ , \underline{b}^+ , \underline{c}^0 . X causes Y and X causes Z. For example, general supervision causes both involvement and production, but involvement does not cause production. X is a causal factor, and Y and Z are resultant factors.

Pattern 6. \underline{a}^+ , \underline{b}^0 , \underline{c}^+ . X causes Z and Y causes Z. For example, both general supervision and involvement cause production. X and Y are causal factors, and Z is a resultant factor.

Pattern 7. \underline{a}^+ , \underline{b}^+ , \underline{c}^+ . X causes Z, X causes Y, and Y causes Z. General supervision can lead to high production in two ways. In one, general supervision influences production by causing involvement which in turn causes production. In the other, supervision causes production directly. In effect, Pattern 7 is a combination of Patterns 1 and 4. X is a causal factor; Y is a causal, resultant, and intervening factor, and Z is a resultant factor. Pattern 7 may also occur when the \underline{c} relationship is spurious, when in fact causal relationships occur as stated in Pattern 5.

Pattern 8. \underline{a}^0 , \underline{b}^0 , \underline{c}^0 . No causal relationships are demonstrated by this pattern. Consequently, we can conclude nothing about the possibility of causal relationships between X, Y, and Z, except that they do not occur with the particular time intervals we chose to measure.

In this discussion of patterns of association, we have specified causal paths moving in only one direction--from X to Y to Z. As Figure 2 indicates, causal paths are also possible in five other directions: (1) X to Z to Y, (2) Y to X to Z, (3) Y to Z to X, (4) Z to X to Y, and (5) Z to Y to X. These paths may be examined in a similar manner.

Patterns which identify Type I and Type II causal relationships.

In order to determine whether Type I or Type II causal relationships occur among X, Y, and Z, it is necessary to compare the causal patterns specified above with portions of causal paths moving in the opposite direction. For example, if we discover that X, Y, and Z are related according to Pattern 1 (\underline{a}^+ , \underline{b}^0 , \underline{c}^0), we must compare the association we obtained between X and Z with the association between Z and X in which Z is measured before X is measured. Such a comparison can lead to one of four conclusions:

1. The association in which X is measured before Z is substantial, while the association in which Z is measured before X is not. A Type Ia causal relationship obtains, and X causes Z. For example, supervision causes production.

2. The association in which X is measured before Z is stronger than the association in which Z is measured before X, and both associations are of substantial size. A Type Ib causal relationship obtains, and X causes Z more strongly than Z causes X. For example, supervision causes production more strongly than production causes supervision.

3. The association in which Z is measured before X is stronger than the association in which X is measured before Z, and both associations are of substantial size. A Type Ib causal relationship obtains, and Z causes X more strongly than X causes Z. For example, production causes supervision more strongly than supervision causes production.

4. The two associations are not substantially different in strength but both are substantial in size. A Type II causal relationship obtains. X causes Z and Z causes X. For example, production causes supervision and supervision causes production.

In a similar manner, we can determine whether Type I or Type II causal relationships obtain in the other specified patterns of association.

EMPIRICAL TEST OF THE METHOD

The use of the method requires measurements of the factors involved in the causal relationship at different points in time, measures of other factors which may affect the causal relationship, and a time lag between measurements which is of a length appropriate to reflect the causal dynamics involved. These conditions were approximated (but not completely fulfilled) in information collected on organizational factors and performance of 151 engineers and from three laboratories of a large electronics firm. These engineers were among 1311 respondents in an extensive study of motivations and working relationships of scientific personnel directed by Dr. Donald C. Pelz. Several consistent associations between organizational factors and job performance were found in that study (they are summarized in Pelz and Andrews, 1966), but it was impossible to determine the direction of causality in these associations.

Several hypotheses were advanced concerning the relationship between organizational factors and output of patents on the basis of previous work in organizational psychology--theory, field studies, laboratory experiments, and a little bit of intuition. Let us consider five hypotheses here in order to illustrate the method for causal analyses.⁹ They are summarized in Figure 4.

Hypothesis 1. Involvement and patents are in a Type Ib causal relationship. Specifically,

- a. Involvement causes patents.
- b. Patents cause involvement.
- c. Involvement causes patents more strongly than patents cause involvement.

According to Hypothesis 1, being involved in his work causes an engineer to

⁹ In the original study twelve hypotheses were tested and four measures of performance were used. For details see Farris (1966) or (1967).

produce more patents, and producing patents causes him to become more involved in his work. The former causal relationship is hypothesized to be stronger than the latter.

Hypothesis 2. Influence and patents are in a Type II causal relationship. Specifically,

- a. Influence causes patents.
- b. Patents cause influence.

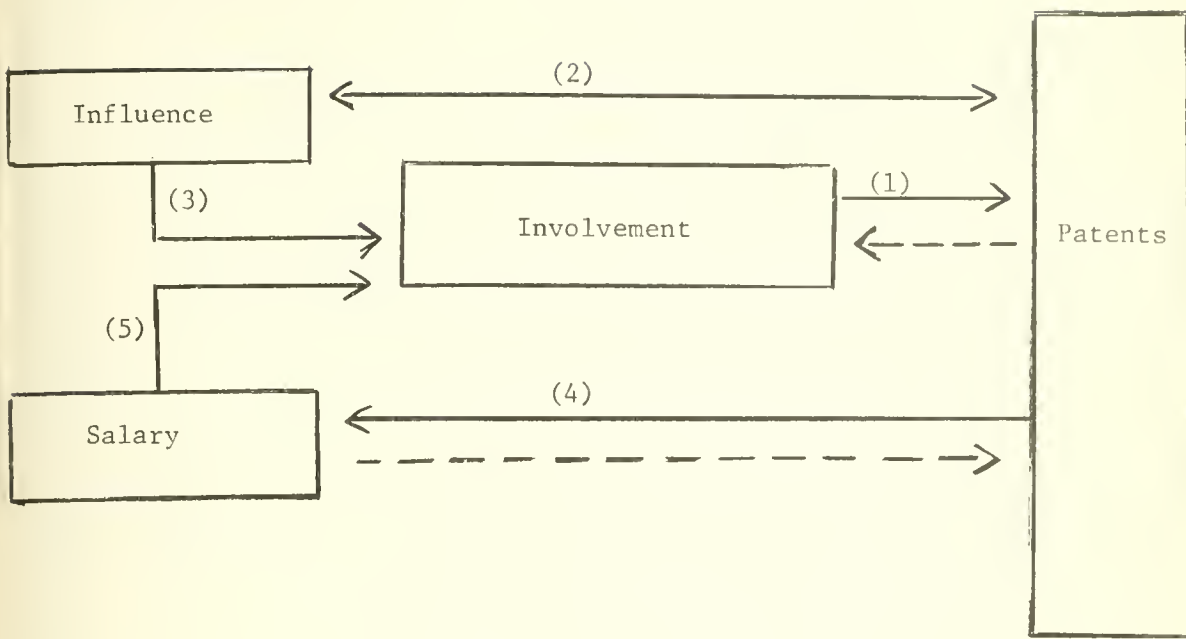
This hypothesis states that influence and patents cause one another, and no significant difference is expected in the strengths of the two causal relationships. Having influence causes the engineer to produce more patents, and producing more patents causes the engineer to have more influence.

Hypothesis 3. Influence and involvement are in a Type Ia causal relationship. Specifically,

- a. Influence causes involvement.

This hypothesis states that involvement is a resultant factor with respect to influence, but no causal relationship is predicted between involvement and subsequent influence.

Thus, influence can lead to patents through either of two paths (see Figure 4). It can affect performance directly (Hypothesis 2) or it can affect involvement (Hypothesis 3) which in turn affects patents (Hypothesis 1). The first path corresponds to Pattern 1 (a^+, b^0, c^0), while the second corresponds to Pattern 4 (a^0, b^+, c^+). The first path is compatible with what Miles (1965) has called the "human resources" approach to management (performance results from utilizing the full capacities of the organization's members), while the second is more characteristic of the "human relations" approach (performance is a function of worker motivation, which is in turn a function of management practices). Together the two paths correspond to Pattern 7 (a^+, b^+, c^+).



- A → B = Type Ia causal relationship (A causes B)
- A $\overset{\text{---}}{\leftarrow}$ B = Type Ib causal relationship (A causes B and B causes A, but A causes B more strongly than B causes A)
- A ↔ B = Type II causal relationship (A causes B and B causes A)

The numbers refer to the five hypotheses.

Figure 4. The predicted causal network.

Hypothesis 4. Patents and salary are in a Type Ib causal relationship. Specifically,

- a. Patents cause salary.
- b. Salary causes patents.
- c. Patents cause salary more than salary causes patents.

Producing patents causes the engineer to get paid more, and getting paid more causes the engineer to produce more patents (the philosophy of merit salary increases). The former causal relationship is hypothesized to be stronger than the latter.

Hypothesis 5. Salary and involvement are in a Type Ia causal relationship. Specifically,

- a. Salary causes involvement.

To the extent that salary serves as a reward and has incentive value, we would expect it to cause variations in the engineer's involvement in his work.

Salary, involvement, and patents are thus hypothesized to be in a relationship similar to that of influence, involvement, and patents. Salary affects performance according to Pattern 1 or Pattern 4, but Pattern 1 is expected to be stronger.

From these hypotheses it is possible to specify certain intervening factors:

1. Involvement is an intervening factor between
 - a. Influence (causal factor with respect to involvement) and patents (resultant factor with respect to involvement).
 - b. Salary (causal factor with respect to involvement) and patents (resultant factor with respect to involvement).
2. Patents is an intervening factor between
 - a. Involvement (causal) and influence (resultant).
 - b. Involvement (causal) and salary (resultant).
 - c. Influence (causal) and involvement (resultant).
 - d. Salary (causal) and involvement (resultant).

3. Influence is an intervening factor between patents (causal) and involvement (resultant).
4. Salary is an intervening factor between patents (causal) and involvement (resultant).

Timing of the measurements.

Self-report questionnaires were received from respondents in 1959 and again in 1965. In each questionnaire the respondent described the organizational factors as he saw them at that moment in time (for example, how involved are you in your work?). The respondent also indicated the number of patents he had produced over the last five years. In 1965 he also reported his output for the last two-and-one-half years. The timing of the measurements is summarized in Figure 5 below.

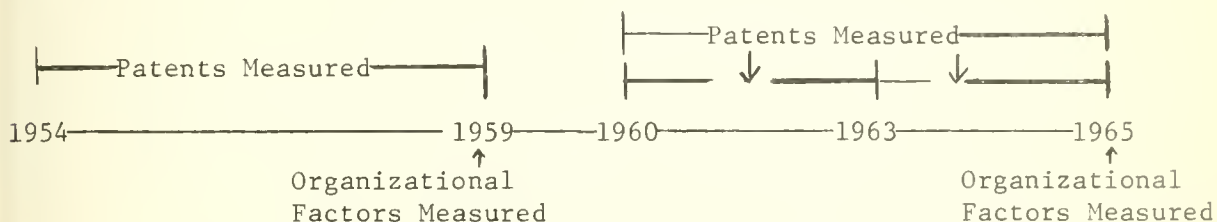


Figure 5. Sequence of data collection in the present study.

Measures of the organizational factors

Involvement. The item measuring involvement asked:

Some individuals are completely involved in their technical work-- absorbed by it night and day. For others, their work is simply one of several interests. How involved do you feel in your work?

CHECK ONE answer. (6-point scale)

Pelz and Andrews (in press, ch. 8) found more consistent relationships to performance with this item alone than with a five-item index which included involvement, interest, identification with task, the importance of his work, and challenge in the scientist's present work.

Influence. The engineer was asked to name the person other than himself who had the most influence on his work goals. Then he was asked to report:

To what extent do you feel you can influence this person or group in his recommendations concerning your technical goals? CHECK ONE.
(5-point scale)

Cases where only the scientist had influence on his work goals were scored as cases of "complete" influence.

Salary. Respondents were asked to indicate their professional income last year from all sources on a 9-point scale.

Measure of patents

One item on the questionnaire asked respondents to report the number of patents they had produced over the past five years. This item was asked in both 1959 and 1965. In addition, a question was included in 1965 asking the respondent to report his output for the last two and a half years. By subtracting responses to this item from those to the previous one, it was possible to determine the respondent's output for the first two and one-half years of the five-year period. Thus, measures of patents were available for the time periods: 1954-59, 1960-65, 1960-62, and 1963-65.

Consistently (e.g., Shockley, 1957; and Pelz and Andrews, 1966), distributions of scientific output have been found to be highly skewed, with most scientists producing very few patents and a few scientists producing many. Skewed distributions of output were found in the present data for the measures of patents over all four time periods. Since interpretation of statistics to be employed in later analyses of the data is made more plausible if distributions on the variables do not depart markedly from normality, the raw output scores were converted to "lognormal" scores. The procedure for doing this was based on a suggestion by Shockley (1957) that distributions of the natural logarithms of numbers of scientific products were reasonable symmetrical and approximated the normal curve. Therefore, in the present study the output score used in each case was $\log_e (\text{raw number of patents} + .5) + 1.0$. The distribution of these scores did not differ much from normality. Details on this adjustment and reasons for including the constants are given in Andrews (1961). This measure of patents was adjusted to hold constant the effects of three background factors which might have led to spurious correlations: high-degree earned, time since degree, and time with laboratory. These adjustments were an attempt to accomplish Step 1 above. For a fuller discussion of the rationale for these procedures, including the reliability and validity of the measures employed, the reader is referred to Pelz and Andrews (1966).

In 1959 about 60% of the respondents received a short-form questionnaire which did not include the questions on influence or salary. Thus, analyses involving 1959 measures of influence on salary are based on an N of about 50, while those involving all other measurements are based on an N of about 125. (The reduction from N = 151 is due to missing data on factors under study or used for adjustment.)

Analysis procedures

Pearson product-moment correlation coefficients were computed among the organizational factors and patents after examining the distributions to make sure that they did not depart markedly from normality. The magnitude and statistical significance of these correlations, determined by one-tailed t-tests, constitute the primary means of testing the hypotheses of this research. Thus, for this study, we have chosen to define "substantially different from zero" in terms of statistical significance. In comparing the sizes of correlation coefficients, we have chosen to examine their relative levels of statistical significance rather than the magnitude of the coefficients themselves. Given different numbers of respondents, sizes of associations, or conditions of measurement, other criteria might have been chosen for defining "substantial."

For our sample size of 50 the appropriate size of the correlation coefficient needed to be significant at the .01 level of confidence is .28. At the .05 level of confidence it is .22. For our sample of 125 the values are .20 and .15 at the .01 and .05 levels, respectively.

Results

For correlations over time to be meaningful, it is necessary to assume that the individual factors being correlated are neither markedly consistent or markedly inconsistent over time (Pelz and Andrews, 1964). Thus, the test-re-test reliabilities of the measures between 1959 and 1965 were determined. They are as follows with the number of cases of parentheses: involvement .46 (133), influence .24 (51), salary .71 (54), patents .39 (130). The 1959 measure of patents correlated .39 with patents for the period 1960-1962 and .27 with patents for the period 1963-1965.

Evidence from another study indicates that the relative instability of these measures reflects changes in the engineer's work situation rather than unreliability in the measuring instruments. Pelz (1962) readministered 89 items from a questionnaire very similar to the one used in the present study to a random sample of 52 scientists two months after they had completed the original questionnaire. Test-retest reliabilities over the two-month period include: involvement .68, influence .66, and patents 1.00. Although salary was not included, its reliability is undoubtedly high.

Table 1 summarizes the data testing the hypotheses. Recall that in each case an attempt has been made to fulfill Step 1 by removing the influence on performance of three factors: highest degree earned, time since degree, and time with laboratory. In each correlation involving patents, patents are measured over the five-year period either immediately preceding or immediately following the measurement of the organizational factors. There is about a six-year time lag between measurements of the organizational factors.

Parts of a and b of Hypothesis 1 are supported by the findings although the stronger relationship tends to be between patents and subsequent involvement. Apparently, being more involved in his work causes the engineer subsequently to produce more patents, but, more than that, patent production causes the engineer subsequently to become more involved in his work.

Hypothesis 2, is supported in part b only. Patents and influence are in a Type Ia causal relationship, patents causing engineers subsequently to have more influence on their work goals. Greater influence on work goals, however, was not shown to increase subsequent performance.

Hypothesis 3 is not confirmed by the data. Although there is a trend for influence to cause involvement, this does not quite reach the .05 level of significance (our definition of "substantial" for this study). There is, however, a significant relationship between involvement and subsequent influence. Although

TABLE 1: Principle Tests of the Hypotheses

	Patents for 5 Years - Immediate
Hypothesis 1. Involvement and Patents	
a. Involvement (X_0) -- Patents (Y_1)	.19*
b. Patents (Y_0) -- Involvement (X_1)	.29**
Hypothesis 2. Influence and Patents	
a. Influence (X_0) -- Patents (Y_1)	.00
b. Patents (Y_0) -- Influence (X_1)	.19*
Hypothesis 3. Influence and Involvement	
a. Influence (X_0) -- Involvement (Y_1)	.21
b. Involvement (Y_0) -- Influence (X_1)	.15*
Hypothesis 4. Patents and Salary	
a. Patents (X_0) -- Salary (Y_1)	.31**
b. Salary (Y_0) -- Patents (X_1)	.01
Hypothesis 5. Salary and Involvement	
a. Salary (X_0) -- Involvement (Y_1)	.24*
b. Involvement (Y_0) -- Salary (X_1)	.20**

* $p \leq .05$ ** $p \leq .01$

TABLE 2. Tests of Hypotheses 1, 2, and 4 Using Different Time Lags

	Patents for 2½ years - immediate	Patents for 2½ years - 2½ year lag	Patents for 5 years - 5 year lag
Hypothesis 1. Involvement and Patents			
a. Involvement (X_0) -- Patents (Y_1)	.18*	.17*	
b. Patents (Y_0) -- Involvement (X_1)	.23**	.27**	.25**
Hypothesis 2. Influence and Patents			
a. Influence (X_0) -- Patents (Y_1)	.15	-.08	
b. Patents (Y_0) -- Influence (X_1)	.17*	.12	.15*
Hypothesis 4. Patents and Salary			
a. Patents (X_0) -- Salary (Y_1)	.20	.33*	.35**
b. Salary (Y_0) -- Patents (X_1)	.13	-.09	

* $p \leq .05$ ** $p \leq .01$

there is a trend for influence to lead to involvement, more than that, being involved in their work seems to cause engineers to have more subsequent influence on their work goals.

Patents and salary (Hypothesis 4) are in a Type Ia causal relationship, more patents causing better pay. There is no support for the idea that better pay causes the engineer to produce more patents.

For salary and involvement (Hypothesis 5) the findings indicate that salary causes involvement as predicted, but involvement also causes salary. The latter relationship reaches a greater level of statistical significance.

Before accepting these findings too hastily, we shall look further to make sure that we have satisfied the conditions needed to establish causal relationships: the correctness of the time lags and the ruling out of the influence of factors other than the two in the causal relationship being examined. Table 2 shows tests of Hypotheses 1, 2, and 4 using different time lags. The first part of the table shows relationships when patents are measured over a two-and one-half year period immediately preceding or following measurement of the organizational factor. For all three hypotheses the findings are in the same direction as they were for the five-year measurement of patents, but in each case the difference between the sizes of the correlations in the contrasting causal directions is reduced considerably. When using the measurement of patents for a 2½-year period and a 2½-year time lag, the original findings for Hypotheses 1 and 4 are again supported. For Hypothesis 2 neither correlation is substantial although the correlation between patents and subsequent influence still tends to be closer to prediction. It was also possible to examine relationships between patents measured over a five-year period and the organizational factors measured five years later. These relationships are again consistent with the original findings in Table 1.

In general, then, the findings apparently hold true for the several time lags and measurement periods used in this study. Although they certainly do not exhaust the universe of time lags, the fact that the findings did occur for them supports the generality of the findings.

Partial correlations were computed to determine whether the findings for Hypotheses 1, 2, and 4 still held for engineers at the same level of past performance or a given organizational factor. Partial correlations between the factors and subsequent patents holding constant past patents were: Involvement .10, influence-.05, salary -.09. Relationships between patents and subsequent amounts of the factors holding constant past amounts of the factors were involvement .15*, influence .13, and salary .29**. Although the partials are generally smaller than the zero-order associations, the findings with them are consistent with those using the zero-order correlations.

The tests of hypotheses 3 and 5 were made by determining the zero-order correlations between one factor measured in 1959 and another measured in 1965. In so long an interval it is possible that events between 1959 and 1965 affected the relationships. The only factor which we were able to measure during that period is patents. Since patents was substantially associated with previous involvement and subsequent influence and salary, partial correlations were computed to determine the effects of involvement on influence and salary through paths not involving intermediate patents. The correlation between involvement and subsequent influence (Hypothesis 3 holding constant patents during the interval) was .12, which is not statistically significant. Apparently, involvement caused patents which, in turn, affected the level of influence. For Hypothesis 5 the correlation between involvement and subsequent salary partialing out patents during the intervening period remained

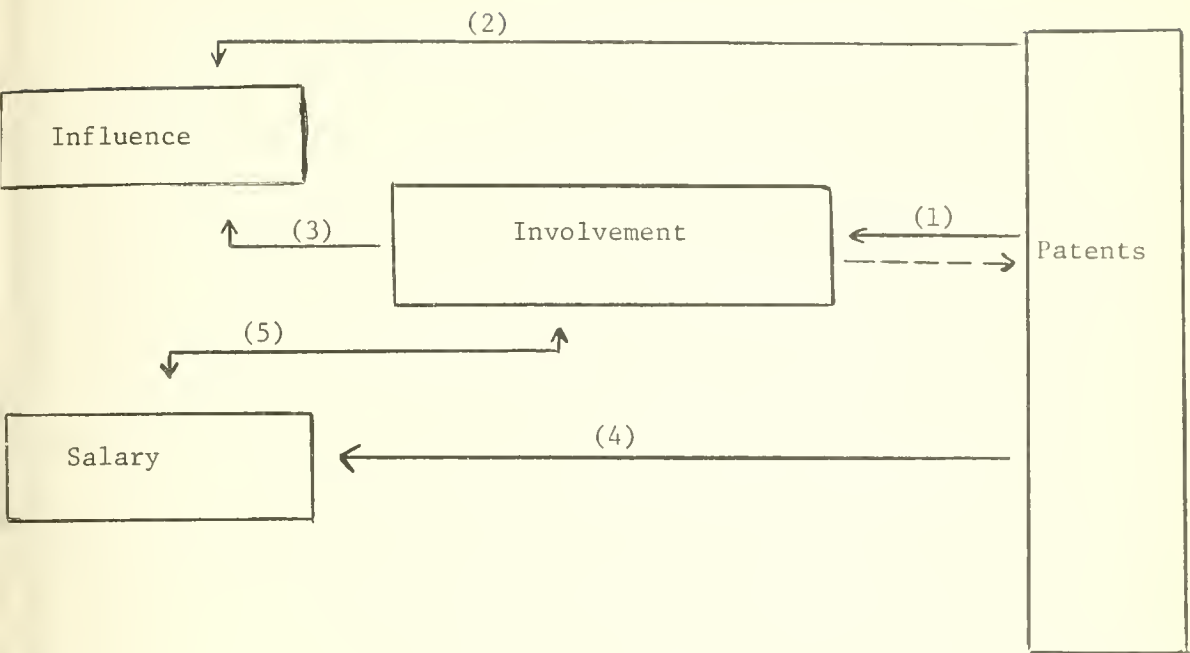
substantial (.16*). Apparently, involvement caused salary according to Pattern 7 through two paths: directly and through patents.

Four other analyses were performed to attempt better to fulfill the definition of a causal relationship. In one eta was used as the measure of association. In another the analysis was repeated for 40 engineers who were in similar job situations throughout the study. These were "bench scientists" who had fewer than four subordinates reporting to them in both 1959 and 1965. In a third attempt the analysis was done separately for each of the three/since aspects of laboratory "climates" may have influenced the causal relationships. In a fourth the analysis was repeated using the absolute number of patents for the 1960-1965 period and log patents for that period unadjusted for the three background factors. With very few inconsistent exceptions the findings from these analyses were the same as those discussed above.¹⁰

In Pelz and Andrews' (1966) original study done in 1959, tests were made of relationships between patents for five years and subsequent involvement, influence, and salary. The results of these earlier analyses are consistent with those of the present study, indicating that these relationships are stable.

The overall findings are summarized in Figure 6, which is based on the analysis using the five-year measure of patents and no time lag, but supported by the other analyses. In brief the findings indicate that patents cause involvement, influence, and salary, and that, of the factors considered, involvement alone causes patents. Salary also causes patents indirectly through its effect on involvement. The following were found to be intervening factors:

¹⁰For details see Farris (1966).



$A \rightarrow B$ = Type Ia causal relationship (A causes B)

$A \overleftarrow{--} B$ = Type Ib causal relationship (A causes B and B causes A, but A causes B more strongly than B causes A)

$A \leftrightarrow B$ = Type II causal relationship (A causes B and B causes A)

The numbers refer to the five hypotheses.

Figure 6. Summary of tentative conclusions.

1. Involvement, between
 - a. Salary and patents
 - b. Patents and salary
 - c. Salary and influence
2. Salary, between patents and involvement.
3. Patents, between involvement and salary.

These findings should be regarded as tentative. They are based on correlations very low in size and possibly unwarranted assumptions that the correlations are not spurious and that the time lags employed actually correspond to the intervals necessary for the alleged causes to exert its influence on the alleged effect. On the other hand, results were consistent using ^{several} time lags (which were all within the broad range suggested by behavioral scientists for organizational factors to affect performance), and the influence of several third factors was diminished by respondent selection and statistical controls. Moreover, in many correlational studies conclusions are based on consistent but small associations, possibly unwarranted assumptions that the associations are not due to third factors, and a time lag (usually zero) which may or may not be the appropriate one for the alleged cause to influence the alleged effect.

DISCUSSION

We have proposed a method for making causal analyses and described an example of its application to social phenomena in the natural setting. Although this application did not fulfill the definition of a causal relationship as much as we would have liked, consistent findings occurred in testing the five hypotheses with different controls for spurious correlation and different time lags. Given the complexity of organizational phenomena, the findings of the empirical study substantiate the usefulness of the method. Now let us contrast the method proposed here with two others: path analysis and panel study analysis. First we shall deal with definitions of a causal relationship used in each method.

Comparison with other methods

Definitions of a causal relationship. Nagel (1961, pp. 73-78) suggested four conditions which a causal explanation should satisfy. "In the first place, the relation is an invariable or uniform one, in the sense that whenever the alleged cause occurs so does the alleged effect. There is, moreover, the common tacit assumption that the cause constitutes both a necessary and a sufficient condition for the occurrence of the effect." Secondly, the events are spatially contiguous. That is, the noon factory whistle in Pittsburgh does not cause workers in New York to go to lunch although they do so immediately after the Pittsburgh whistle blows. Thirdly, the alleged cause precedes the alleged effect in time and is also continuous with it. Finally, the relation is asymmetrical. That is, general supervision causes performance, but nothing is said about high performance causing general supervision.

Let us examine the definitions of a causal relationship proposed in Lazarsfeld (1946), and Simon (1957), and the present study in terms of these four conditions. To examine these particular definitions is especially useful, since the author of

each made a major contribution to the study of causal relationships in the behavioral sciences. Lazarsfeld was an early proponent of causal analyses of panel data, and Simon's work with path analysis has stimulated a considerable amount of work in this area.

Lazarsfeld (1946, pp. 124-24) states:

We can suggest a clearcut definition of the causal relationship between two attributes. If we have a relationship between "x" and "y"; and if for any antecedent test factor the partial relationships between x and y do not disappear, then the original relationship should be called a causal one. It makes no difference whether the necessary operations are actually carried through or made plausible by general reasoning.

Simon (1957, pp. 10-35, 50-61) defines causality and "causal relation" in the language of symbolic logic. At the risk of doing injustice to his extensive treatment of the problem of defining causality, let us consider this statement of his: (1957, pp. 34-35)

Causality is an asymmetrical relation among certain variables, or subsets of variables, in a self-contained structure. There is no necessary connection between the asymmetry of this relation and asymmetry in time, although an analysis of the causal structure of dynamical systems in econometrics and physics will show that lagged relations can generally be interpreted as causal relations.

1. Invariability. All three definitions begin with an association between two variables which does not have to be a perfect one. (Although Simon treats only invariable cases in defining, he and others who use path analysis make extensive use of statistical relationships. See, for example, Simon (1954) and Blalock (1961a).

2. Spatial contiguity. None of the definitions discusses this point explicitly although all assume it. Simon's point of a "self-contained structure" may imply spatial contiguity for him.

3. Time sequence. It is the keystone of the definition proposed here. Lazarsfeld says nothing at all about the sequence of the associated variables. Simon specifically excludes time sequence from his definition but notes that it may well be a characteristic of a causal relation.

4. Asymmetry. It is the central aspect of Simon's definition. Although he allows for symmetrical causal relationships, Lazarsfeld fails to mention this condition. The definition proposed here is asymmetrical in its defining operations. X measured at Time 1 causes Y measured at Time 2, but Y measured at Time 2 does not cause X measured at Time 1.

In sum, Lazarsfeld's definition fails to meet any of Nagel's conditions except spatial contiguity (implicitly) and invariability (statistically). Simon's satisfies all four except that of time sequence, with which it does not conflict. If we accept Nagel's notion of causality as a criterion, then the definition advanced here is as good as Simon's and better than Lazarsfeld's. All three treat invariability statistically. Based upon asymmetry, Simon's definition does not specify time sequence but is compatible with this notion. Based upon time sequence the present definition is compatible with the notions of both asymmetrical and symmetrical relationships. Based on neither asymmetry nor time sequence, Lazarsfeld's definition identifies causal relationships but does not determine which of the associated factors is the cause and which is the effect.

Application of the methods

Path analysis specifies the patterns of influence among factors measured at one point in time. A weight or "path coefficient" is assigned to each factor on the basis of a partitioning of the original association into component parts or paths. Resembling aspects of regression analysis, this partitioning utilizes patterns of association similar to those described in this paper. Reports

on path analysis may be found in Wright (1921, 1934, 1960a, 1960b), Li (1955, 1956), Tukey (1954), Turner and Stevens (1959), and Duncan (1966).

In Simon's (1954, 1957) approach models are established which predict different causal relationships among factors. These models are expressed in the form of simultaneous equations which, when solved, yield values of the possible path coefficients. The model which comes closest to describing the obtained data is selected as the appropriate one. Following Simon, Blalock (1960, 1961a, 1961b, 1962a, 1962b, 1964) has developed detailed models for causal analyses in cases where data are available on four or five factors.

Use of the panel study to make causal analyses has been advocated by Lazarsfeld (1954) and his associates (Lipset, et al., 1954; Lazarsfeld and Rosenberg, 1955; Kendall and Lazarsfeld, 1950; Hyman, 1955, ch. 7). In the panel study the same measurements are taken on the same people at two or more different points in time. To determine which of two factors caused the other, comparisons analogous to those proposed in this chapter for distinguishing between Type I and Type II associations are made. The association between the first factor measured at Time 1 and the second measured at Time 2 is compared with the association between the first measured at Time 2 and the second measured at Time 1. The factor measured at Time 1 in the larger association is the probable causal factor. Lazarsfeld's particular method involves a 16-fold table which displays frequencies for each of two factors measured on two occasions.

Campbell (1963), Campbell and Stanley (1963), and Pelz and Andrews (1964) have extended the logic of the panel study of dichotomous factors to situations in which continuous factors are available. Correlation coefficients between two factors measured on two occasions are compared. Campbell has suggested calling this approach the method of "cross-lagged panel correlation."

In examining the relationships among path analysis, cross-lagged panel correlation, and the present framework it is helpful to think in terms of the two questions the analysis purports to answer: (1) is a given factor causal, intervening, or resultant with respect to other factors? and (2) is it in a Type I or Type II causal relationship with each of the other factors with which it is associated? In answering the first question, the framework proposed here resembles path analysis except for two things: the factors are measured at different points in time, and patterns of association rather than the path coefficients are used to determine the answer. In answering the second, the framework employs the same logic as cross-lagged panel correlation analysis. By measuring factors at more than one point in time, the present framework allows path-analytic equations (rather, patterns of association) to be started with greater certainty. By treating several factors at the same time and proposing a scheme for causal analysis among factors all of which are not measured at the same two points in time, it overcomes some limitations of cross-lagged panel correlation in its current formulations.

Assessment of the present framework.

We have seen that the definition of a causal relationship proposed here does no more injustice to the four common philosophical aspects of causality than either the path analytic or panel-study approach. In fact it incorporates some important characteristics of each. All three approaches have distinct advantages in studying social phenomena in the natural setting in that they perturb the system under investigation much less than an experimental study and, unlike the correlational study, they allow conclusions about causality to be drawn. The ease with which the notion of causality presented here can be made operational is demonstrated by the fact that research designed

around it allowed us to test hypotheses about social phenomena in the natural setting based on descriptions of other behavioral scientists.

Two aspects of the framework presented here appear to be critical in its usefulness and provide it with distinct advantages over path analysis or panel studies in their current formulations. First, unlike the common applications of the other methods, it allows for symmetrical causal relationships. A definition of causality which permits only asymmetry does not allow conclusions to be drawn that, for example, involvement causes patents and patents cause involvement. Yet undoubtedly causality in ongoing social systems is often a two-way street. The distinction between escalation and response to enemy escalation is a fine one. A problem with many experimental designs is that, based on the implicit assumption of causality as symmetrical, they fail to examine the causal hypothesis opposite to their predictions. Moreover, they consider few of the several relationships undoubtedly involved in the complex causal cycles of social phenomena.

A second critical aspect is that it forces us to study time lags between cause and effect. It is probable that in the natural setting cycles can be charted which show, for example, how involvement changes over time and how patents change over time. Depending on the shape of these curves, a positive, negative, or zero association may occur between them at a single point in time or at a given time lag. In detecting causal relationships, it is important that the factors be measured at intervals corresponding to the time lag needed for one factor to affect the other. Thus, there is nothing sacred or methodologically pure about the one-point-in-time correlation! It is only one of an infinite number of possible time lags over which the factors may be measured, and there are no data available to show that measuring both factors at the same

point in time is more apt to capture the true nature or strength of their causal relationship. Definitions of causality emphasizing simultaneity of cause and effect (e.g. Lewin, 1942, 1943) fail to cope fully with this point although their argument that some "trace" of the alleged cause should be present at the time the effect occurs is a valid one. Examination of test-retest reliabilities is one way of looking for such a trace.

Confident statements about causal relationships among social phenomena in the natural setting can be made only through a method of research which examines relationships over time. The present framework offers one alternative for doing so. Although it is prone to error by mis-estimating time lags, it is no more so than single-point-in-time correlational studies, and it does allow investigation of time lag phenomena. Like many other methods using correlations, it allows conclusions to be drawn with greater certainty when correlation coefficients are relatively large and third factors are controlled; however, its application to the present study was quite successful, despite the low-magnitude correlations involved and the lingering possibility of spurious correlations.

Well-designed experiments should be performed wherever possible. However, it is difficult to manipulate social phenomena in the natural setting, and experiments to date have not examined the reverse causal hypothesis--that the alleged dependent factor causes the alleged independent factor. Single-point-in-time correlational studies are feasible in the natural setting, but they do not allow conclusions to be drawn about causal relationships with any ease or precision. It may be wise to consider a three-phased approach to the study of social phenomena in the natural setting: (1) correlational studies to determine generally what is associated with what, (2) use of a method like the present

one to determine generally the cycles of causal relationships involved, and
(3) experiments to determine more precisely the causal relationships between
phenomena of particular interest.

REFERENCES

- Andrews, Frank M., 1961. "Logarithmic transformation of output of scientific products." Study of Scientific Personnel Analysis Memo No. 11, pp. 163-170. Ann Arbor: Survey Research Center, Institute for Social Research, The University of Michigan.
- Blalock, H. M., 1961. Correlational analysis and causal inferences. Amer. Anthropologist, 62:624-631.
- Blalock, H. M., 1961a. Correlation and causality: the multivariate case. Social Forces, 39:246-251.
- Blalock, H. M., 1961b. Evaluating the relative importance of variables. Amer. Sociological Review, 26:866-874.
- Blalock, H. M., 1962a. Spuriousness versus intervening variables: the problem of temporal sequence. Social Forces, 40:330-336.
- Blalock, H. M., 1962b. Four variable causal models and partial correlations. Amer. J. Sociology, 68:182-194.
- Blalock, H. M., 1964. Causal inferences in nonexperimental research. Chapel Hill: University of North Carolina Press.
- Campbell, D. T., 1963. From description to experimentation: interpreting trends as quasi-experiments. in Chester Harris (ed.), Problems in measuring change, pp. 212-242. Madison: University of Wisconsin Press.
- Campbell, D. T., and Stanley, J. C., 1963. Experimental and quasi-experimental designs for research on teaching. in Nathaniel L. Gage (ed.), Handbook of research on teaching, pp. 171-246. New York: Rand McNally.
- Duncan, O. D., 1966. Path Analysis: Sociological Examples. American Journal of Sociology, 1966, v. 72, pp. 1-16.
- Farris, George F., 1966, A causal analysis of scientific performance, Doctoral dissertation, University of Michigan, Ann Arbor, University Microfilms No. 66-14,517.
- Farris, George F., 1966, A longitudinal study of scientific performance. M.I.T. Sloan School of Management Working Paper No.
- Hyman, Herbert., 1955. Survey design and analysis. Glencoe, Ill: Free Press.
- Kendall, P. L., and Lazarsfeld, P. F., 1950. Problems of survey analysis. R. K. Merton and P. F. Lazarsfeld (eds.), Continuities in social research. Glencoe, Ill: Free Press.
- Lazarsfeld, Paul F., 1946. "Interpretation of statistical relations as a research operation." Address given at the Cleveland meeting of the American Sociological Society. Reprinted in P. F. Lazarsfeld and M. Rosenberg, The language of social research. Glencoe, Ill: Free Press, 1955.

- Lazarsfeld, P. F., 1954. "Mutual effects of statistical variables." Paper read at Dartmouth seminar on social processes, Hanover, N. H., July.
- Lazarsfeld, P. F., and Rosenberg, Morris. (eds.) 1955. The language of social research. Glencoe, Ill: Free Press.
- Lewin, Kurt, 1942. Field theory and learning. The psychology of learning, ch. 4. Natl. Soc. Stud. Educ. 41st yearbook, Part II, pp. 215-242.
- Lewin K., 1943. Defining the "field at a given time." Psychol. Review, 50:288-290; 292-310.
- Li, C. C., 1955. Population genetics, ch. 12. Chicago: University of Chicago Press.
- Li, C. C., 1956. The concept of path coefficient and its impact on population genetics. Biometrics, 12:190-210.
- Lipset, S. M., Lazarsfeld, P. F., Barton, A. H., and Linz, J., 1954. The psychology of voting: An analysis of political behavior. in Garner Library (ed.), Handbook of social psychology, pp. 1124-1175. Cambridge, Mass: Addison Wesley.
- Miles, R. E., 1965. Human relations or human resources? Harvard Business Review, July-August, pp. 148-163.
- Nagel, E., 1961. The structure of science. New York: Harcourt, Brace, and World.
- Pelz, D. C., 1962. Reliability of selected questionnaire items. Preliminary Report No. 9, Study of Scientific Personnel. Ann Arbor: Survey Research Center, Institute for Social Research, The University of Michigan, August.
- Pelz, D. C., and Andrews, F. M., 1964. Detecting causal priorities in panel study data. Amer. Sociological Review, 29:836-848.
- Pelz, D. C., and Andrews, F. M., 1966. Scientists in organizations. New York: Wiley.
- Shockley, W. 1957. On the statistics of individual variations of productivity in research laboratories. Proceedings of the IRE, 45:279-290.
- Simon, H. A., 1954. Spurious correlation: A causal interpretation. J. of the Amer. Statistical Association, 49:467-479.
- Simon, H. A., 1957. Models of man. New York: Wiley.
- Tukey, J. W., 1954. Causation, regression, and path analysis. in Oscar Kempthorne (ed.), Statistics and mathematical biology. Ames: Iowa State College Press.

- Turner, M. E., and Stevens, C. D., 1959. The regression analysis of causal paths. Biometrics, 15:236-258.
- Wright, S., 1921. Correlation and causation. J. of Agricultural Research, 20:557-585.
- Wright, S., 1934. The method of path coefficients. Annals of Mathematical Statistics, 5:161-215.
- Wright, S., 1960a. Path coefficients and path regressions: Alternative of complementary concepts? Biometrics, 16:189-202.
- Wright, S., 1960b. The treatment of reciprocal interaction, with or without lag, in path analysis. Biometrics, 16:423-445.

