The Journal of Mind and Behavior
Winter 1991, Volume 12, Number 1
Pages 49–70
ISSN 0271-0137

Inferring Formal Causation from Corresponding Regressions

William V. Chambers

University of South Florida

A statistical method for inference of formal causes was introduced. The procedure, referred to as the method of corresponding regressions, was explained and illustrated using a variety of simulated causal models. The method reflects IV/DV relations among variables traditionally limited to correlational or structural equation analysis. The method was applied to additive, subtractive, multiplicative, recursive and reflected models, as well as models of unrelated and correlated dependent variables. Initial applications to data from physical science, biology, economics, marketing and psychology were developed, with generally supportive results.

The development of teleological theories has been impeded by a shortage of suitable statistical methodologies. Empirical researchers tend to confound theoretical models of causation with methods of inferring independent variables (IV) and dependent variables (DV) [see Rychlak, 1981, pp. 57–60]. Since the traditional methods of empirical research are based on mechanistic conceptions of causation, there has been a tendency for telic theories to be reduced to non-telic, billiard ball type models, when elaborated by traditional statistics. The present paper proposes a statistical methodology for inference of non-mechanistic IV/DV relationships, thus making the match between theory and method more consistent for the teleologist.

A full appreciation of the mismatch that has occurred between theory and method requires a review of causal terminology (ibid, pp. 120–123). Traditional causal theory stems from Aristotle, who suggested there are at least four types of causation: material, efficient, formal and final. These four types of causation are best illustrated by example. A chair may be explained as the product of all four of Aristotle's causes. It is a chair because it is made of wood, as opposed to jelly (material cause). It is a chair because a craftsman

Requests for reprints should be sent to William Chambers, Ph.D., Department of Psychology, University of South Florida, Fort Myers, Florida 33919.

sawed, sanded, glued and varnished the wood during construction (efficient cause). It is a chair because of its shape, which consists of a seat, a back, two arms and four legs, all of which have structural relationships with one another (formal cause). Finally, it is a chair because someone intended it to be one for a reason (need, desire, wish, etc.). Each of these perspectives on causation clearly contributes to an understanding of the chair. There is little reason to believe these four perspectives should not similarly contribute to an understanding of a person. Rychlak (1981) points out, however, that psychologists have been overly dependent on efficient cause theories, primarily out of reliance on statistical methods developed from efficient cause perspectives. Although the abandonment of methods altogether would end the tendency of psychologists to confound theory with methods, both theory and method are essential in scientific endeavors. Indeed, they can facilitate one another's growth, as long as the match is appropriate. The primary goal of this paper is to introduce a formal cause statistical method that matches the requirements of formal cause theorization.

Although elegant procedures for inference of efficient causation have been developed, as best illustrated in the work of Cook and Campbell (1979), these methods have not been particularly successful when applied to formal cause studies. Correlational studies have generally been used to reflect the structures of relevance to formal causes. Since random assignment and physical manipulation of variables are rarely used in correlational studies, however, causal inferences from correlations are largely speculative. The crucial element of time, which is at the core of efficient cause inferences, is absent in most correlational studies. The absence of the time dimension in correlational studies is, however, not simply a matter of inconvenience. Many constructs addressed in correlational studies do not incorporate time at any level, not even at theoretical levels, but make sense only as expressions of formal causes. The investigator interested in formal causes cannot simply say "the sawing caused the chair because it preceded the chair." From a formal cause perspective, the chair is a chair due to the simultaneous coordination of a back, seat, arms and legs. Creativity may be similarly construed as the simultaneous synthesis of divergent thinking and coherence. In a sense, the notion of time may not be relevant to an understanding of either creativity or a chair. A chair would be a meaningful construction even if it had always existed. The use of correlations alone to study such atemporal models, unfortunately, exacts a heavy price for formal cause theoreticians.

The most serious problem with the use of correlations in formal cause studies is the failure of correlations to reflect the hierarchical constructs used in most theories. Although correlations do offer some benefits in the assessment of the structure of complex constructs like creativity, they do not reflect implicit structural asymmetries. The correlation of creativity with coherence is the same as the correlation of coherence with creativity. Correlations cannot disclose the dependency of creativity on divergent thinking and coherence, nor the independence of divergent thinking and coherence from one another and creativity. Without the development of procedures to assess such asymmetrical implications, psychology has failed and will continue to fail to empirically validate theories based on formal causes. As will be illustrated later, this has serious implications for all scientists but especially for those interested in teleölogical theories.

In spite of the difficulty of the inference of asymmetrical relations, the quest to do so is omnipresent in science. Simon (1977) has argued that most modern scientists implicitly assume that causality consists of asymmetrical functional relations. The clearest expression of such asymmetrical relations is probably $x \Rightarrow y$; that is, x causes y, but y does not cause x. In essense, this term implies variable y is dependent on variable x, while x is independent of y. The development of structural equations and path analysis have made it possible to use covariance structures to assess the plausibility of hypothesized causal models, with the direction of cause being left to theoretical assumptions outside of the data. James, Muliak, and Brett (1982) provide an introduction to causal modeling and the difficulty of gleaning the existence of asymmetrical relations from actual data. In spite of the elegance of the mathematics involved in structural equations, the problem of asymmetry has not been solved by structural equation methods. Consequently, the measurement of what most people think of as causality, i.e., asymmetrical relations, is still largely a speculative endeavor. In a previous study, however, the present author (Chambers, 1986) described a statistical method, referred to as the method of corresponding variances, which provides an assessment of asymmetrical relations. The procedure to be developed in the present study is an extension of the principle of corresponding variances and is referred to as the *method of corresponding regressions*. The new method is assumed to be a potentially powerful alternative or addition to structural equation techniques in that it makes possible the inference of asymmetrical relations independent of time, and consequently, facilitates the elaboration of formal cause theories.

In the previous study the author (Chambers, 1986) suggested that restricting the range of an independent variable should restrict the variance of corresponding values of a dependent variable. When a dependent variable is generated by the summation of two independent variables, the extremely high values of the dependent variable are likely to arise from combinations of higher values of the independent variables while low values of the dependent variable are likely to arise from low values of the independent variables. Moderate values of the dependent variable may arise from many levels of the independent variables, since opposite values of the independent variables may combine to cancel to the mean of the dependent variable. Moderate values in the independent variables, however, limit the range of values in the dependent variable. A moderate independent value is likely to combine with another independent value to create a moderate dependent value. This follows from the assumption that the dependent variable is a summation of the independent variables and that the independent variables thus directly determine the dependent variable. In the previous study, simulations were used to demonstrate that restricting the dependent values to those corresponding to the mid-range of the independent variables did restrict the variance of the dependent values when compared to the variance of independent values corresponding to the mid-range of the dependent variable. The conclusion of the study was that this asymmetry of variances within corresponding mid-ranges could be useful in the inference of formal causal relations in certain contexts.

The present study involves a closer look at the tendency for higher values of independent variables to combine to create higher dependent values while lower independent values combine to create lower dependent values. The causal model used in the previous study will be used again in the present study. Two uniformly distributed orthogonal random variables, $x1$ and $x2$, each with one hundred observations ranging from zero to one were generated via computer. Uniform variables were used in order to simulate uniform relations between continuous variables, allowing for general statements of the relations between the variables across their ranges. A dependent variable, y, was simulated by adding the x variables, i.e., $y = x1 + x2$. Thus a model of asymmetrical relations, symbolized as $x => y$, was generated, with y being determined by the values of $x1$ and $x2$. Furthermore, since y is tautological with $x1$ and $x2$ (independent of time), the configuration is assumed to model a type of formal causation.

Asymmetrical Implications and Corresponding Regressions

In the model $y = x1 + x2$, higher values of y are likely to arise from combinations of high $x1$ and high $x2$ values while lower vaues of y are likely to come from low $x1$ and low $x2$ values. Thus the extreme values of y are likely to correspond to those incidents when similar $x1$ and $x2$ values are added. Mid-range y values, on the other hand, are likely to emerge when dissimilar x values cancel one another toward the mean of y. In other words, $x1$ and $x2$ should be positively correlated at the extremes of y and negatively correlated at the mid-range of y. No such polarization should exist, however, between the x variables across the ranges of x variables, since neither x variable implies the other. Because $x1$ and $x2$ are orthogonal, the expected correlation between the two variables is zero, whether restricting the correlations to the extremes and mid-ranges of x values or not. The polarization of correlations among *x* variables across the ranges of *y*, but not across the ranges of the *x* values themselves, is *the kernel* of the method or corresponding regressions.

The above assumptions were tested by examining the correlations between *x1* and *x2* at the extreme (upper and lower 25 ranks) of *y* and at the mid-range (middle 50 ranks) of *y*. The same simulation described above was replicated 50 times, with *y* being generated by adding orthogonal *x1* and *x2* variables. The average correlation between *x1* and *x2* at the extremes of *y* was $r = .57$. The average correlation between *x1* and *x2* at the mid-range of *y* was negative, $r = .77$. The *x* variable correlations did not tend to polarize when sorted by *x*. The mean correlations between *x1* and *x2* at the extremes and mid-ranges of *x1* were, respectively, $r = .04$ and $r = .02$. Similar findings transpired when the sorting followed the ranks of *x2*. This simulation thus supports the tenability of the proposed kernal of the method of corresponding regressions.

The division of data into extremes and mid-ranges is somewhat arbitrary and clumsy. The method of corresponding regressions, on the other hand, will be accomplished by using deviation scores and regression analysis, leading to a more elegant and efficient solution than that presented in the method of corresponding variances. The new procedure requires conducting two regression analyses on the same two variables, letting each serve as the predictor of the other in turn. First, either an *x* or *y* is used to predict the other, with the understanding that in eventual applied research their status as IVs and DVs will not be known. As the regression models are developed, the prediction errors are derived by subtracting the predicted values from the actual values of the predicted variable. These errors of prediction are then converted to absolute values in order to reflect the extremity of the errors. Next, the absolute values of the deviations from the mean of the predictor variable are determined in order to reflect the extremity of the predictor values. The correlation between these absolute deviations and the absolute errors is found. When the predictor variable is *x* this correlation is symbolized as *rde*(*x*) – the correlation between absolute deviations from the mean of *x* and the absolute errors from predicting *y*. When *y* is the predictor variable this index is referred to as *rde*(*y*) – the correlation between absolute deviations from the mean of *y* and the absolute errors from predicting *x*. The assessment of asymmetrical relations comes from a comparison of *rde*(*x*) and *rde*(*y*). The *rde*(*y*) value should be more negative when the real dependent variable serves as the predictor.

On the surface these simulations and procedures may seem impractical in that most analyses of real data will not explicitly incorporate the equivalents of *x1*, *x2*, and *y*. At best a researcher might hope to measure only part of the independent varuables that determine a *y* variable. In the simplest case there might be only two variables measured, *x1* and *y*. Since the procedure described above requires three variables, it might seem that the method is useless. Fortunately there is a solution to this ostensible problem of model underidentification.

Since *y* is the sum of *x1* and *x2*, errors in the prediction of *y* from *x1* will be perfectly correlated with *x2*. Similarly, errors in the prediction of *x1* from *y* will also be perfectly correlated with *x2*. Consequently the regression analysis residuals can be used to reflect variables that are part of the causal model but that have not been explicitly measured and incorporated in the experimental design. The error term of course becomes *x1* when *x2* and *y* are taken as the predictor and predicted variables. Without this convenient relationship between errors and the components of the causal model, the method of corresponding regressions would not be practical. That it exists provides an avenue for examining asymmetrical relations with ease.

Consider the symmetry of the relationships between *x1* and *x2* (error) across the levels of *x1* and *y*. When predicting *y* from *x1* the relationship between *x1* and error (*x2*) is uniform, that is, there is no polarization of the relationships between *x1* and *x2* across the extremes of *x1*. However, since the relationships between *x1* and *x2* determine jointly the value of *y*, the relationship between error (*x2*) and *x1* is not uniform across the levels of *y*. At the extremes of *y*, *x1* and *x2* correlate one way, at the mid-range, another way. Since least squares regression tends to minimize prediction errors, the question arises as to which tendency will dominate to form the regression slope when predicting *x1* from *y*, the correlation between *x1* and *x2* (error) at the extremes of *y* or the opposing correlation at the mid-range of *y*?

The answer follows from the logic of least squares regression. The prediction slope when predicting *x1* from *y* will tend to be oriented by the extremes of *y*. An extreme value can lead to a relatively extreme error, especially since least squares analysis weights errors by squaring them. Thus, in order to reduce the overall error, the extremes of *y* will orient the prediction slope when predicting *x1* from *y*. Consequently, absolute errors predicting *x1* from *y* should correlate negatively with the absolute values of the deviations of the *y* values from their mean, since absolute deviations reflect extremity. Error will not be asymmetrically reduced across the levels of *x1* when *x1* serves as the predictor of *y*. This is because the correlation between *x1* and error (*x2*) remains uniform across both extremes and mid-ranges of *x* variables. Therefore, *rde*(*y*), the correlation of the absolute errors predicting *x1* from *y*, should be more negative than *rde*(*x*), the correlation between absolute deviations from *x1* and absolute errors predicting *y*. This difference in the correlations of predictor extremities and regression residuals is a reflection of the asymmetrical formal relationships between IVs and DVs and is the basis of the method of corresponding regressions.

IV/DV Correlation, Sample Size and Power

Limitations of Basic Model

The model of causation that has been proposed is rather simple. It was suggested that the model $x => y$ can be simulated by adding uniform random x variables to create dependent y values. Although the model may have immediate intuitive appeal, many questions arise as to its usefulness. A number of side conditions are implicit in the model and these restrictions could effect the utility of the method of corresponding regressions, particularly to real data. A partial, although non-exhaustive, list of potential problems arises from assumptions that the IV and DV correlate $r = .7$ and that nature has developed an additive rather than multiplicative or subtractive model. There is also the problem of correlated dependent variables to address, as well as the ubiquitous problem of sample size. Some attempt will be made to address each of these problems in the following pages, although the reader should not expect the last word on any of these issues for quite sometime – either in favor or against the method of corresponding regressions. The goal for the present is simply to initiate research concerning limitations of the method, leaving more detailed and conclusive judgements to future research and experience.

Level of Correlation

In the proposed causal model, the simulated DV and IV should typically correlate $r = .70$. This follows from the creation of y by adding $x1$ and $x2$. Each IV explains about 50% of the variance of y. Because the method of corresponding regressions pivots upon the relationship between $x1$ and $x2$, where $x2$ is equated with error in predicting y and $x1$, the efficiency of the method is likely to decrease as the data deviate from the simple model. To the extent x and y fail to correlate $r = .7$, the correlation between the predictor extremities and the absolute prediction errors should decrease. As the correlation of the y values with the $x1$ values decreases, the prediction of the variables from one another decreases, and the correspondence that brings about the correlation between x values at the extremes of y also decreases. On the other hand, if $x1$ and y become so highly correlated that there is little error, there can be little to gain from assessment of how the errors differ in the corresponding regressions. In summary, the method should be less efficient when the x and y values deviate from the .7 correlation, since the extremity of y will simply be less relevant to the relationship between $x1$ and $x2$ (error).

To explore the effect of varying the correlation between IVs and DVs, several versions of the causal equation were developed. Three conditions were created by giving different weights to the *x1* values before generating the *y* values. The weights altered the size of *x1*. In the first condition *x1* and *x2* were weighted by being multiplied by the number 1. Each *x* variable thus contributed equally to the determination of *y* in the first condition. Fifty replications of this model, with $n = 100$, were conducted. A summary statistic, *D*, was found by subtracting *rde*(*x*) from *rde*(*y*). In the present and all subsequent simulations, the *x1* variable was used as the *x* variable in deriving *D*. Previously it was hypothesized that *rde*(*y*) would be more negative than *rde*(*x*). The expected value of *D* is thus a negative value. In the first condition, in which *x1* and *x2* were weighted equally by 1, the average correlation between *x1* and *y* was $r = .72$. The average *rde*(*y*) was $-.48$. The average *rde*(*x*) was $-.01$. The average *D* was thus $-.47$, with the $sd = .11$. A *t*-test based on these differences produced $t = -31.34$, $p < .0001$. Thus, as expected, the *rde* correlation was more strongly negative (as opposed to zero as suggested by the null hypothesis) when the dependent variable was the predictor.

A second condition was developed by multiplying *x1* by .25 in order to lower the correlation between *x1* and *y*. In this and all subsequent conditions *x2* was weighted by 1. The average correlation between *x1* and *y* across the fifty simulations of this condition was $r = .21$. The average *D* under this condition was $-.08$, $sd = .07$, $t = -7.36$, $p < .0001$.

A third condition was created by weighting *x1* to correlate above .7 with *y*. To do this *x1* was multiplied by 2 before the generation of *y*. As in the above conditions, *x2* continued to be multiplied by 1. The average correlation between *x1* and *y* in this condition was $r = .90$. *The average rde*(*y*) was $-.31$ and the average *rde*(*x*) was $-.03$. The average *D* for this third condition was $-.28$, $sd = .10$, $t = -19.23$, $p < .0001$. The value of *D* thus predicted the IV/DV variables under this level of correlation as well.

In summary, the average *D* values for conditions one, two, and three were respectively -31.34, -7.36, and -19.23. Casual inspection of the means suggests that correlations between *x1* and *y* that deviate from $r = .7$ reduce the efficiency of the method of corresponding regressions. The fact that substantial *t* values were found in each condition suggests that, at least for large data sets, i.e., $n = 100$, the method may be useful in inferring asymmetrical relations. The values of *rde*(*y*) did tend to be negative, the values *rde*(*x*) did hover about zero and the values of *D* did tend also to be negative, rather than zero, as would be the case if the method did not work. The power of *D* does appear, however, to be affected by the value of the correlation between *x1* and *y*. Although the above statistics argue that the method detected asymmetrical relations under different levels of correlation between *x1* and *y*, no attempt was made to formally assess trends across conditions. This is because it was expected that such analysis would ignore the impact of varying the sample size per replication. There seems little reason to doubt that sample size will play a role in the power of the D statistic. Consequently, a detailed consideration of power will be postponed briefly, in order that the problem of varying sample sizes may be first addressed.

Sample Size

As in most statistical procedures, greater reliability and precision are likely to be found when corresponding regressions are conducted on larger data sets.To initiate an inquiry into the relevance of sample size to the efficiency of the method of corresponding regressions, three conditions were developed, using the previous scheme, weighting both $x1$ and $x2$ by 1. As before, fifty replications per condition were conducted. In the first condition n was set at 100. As in the previous simulations the $rde(y)$ was, on average, strongly negative, $rde(y) = -.48$, $rde(x)$ near zero; $rde(x) = 0.00$, while D tended to be negative, $D = -.48$, $sd = .12$, $t = -28.35$, $p < .0001$. The average correlation between $x1$ and y in this condition was $r = .71$. The second condition was based on fifty replications of $n = 50$. The mean $rde(y)$ was $-.44$, the mean $rde(x)$ was $-.02$ and the mean D was $-.42$, $sd = .20$, $t = -11.15$, $p < .0001$. In the third condition there were fifty replications with $n = 25$. The mean $rde(y)$ was $-.50$, the mean $rde(x)$ was $-.06$ and the mean D was $-.44$, $sd = .25$, $t = -11.26$. Thus in all of the conditions the value of D tended to be similar and significantly negative. Casual inspection of the standard deviations of D suggests, however, that the power of D varies across conditions. Again, no formal analysis of trends was conducted because it was assumed that a more meaningful power analysis would occur by examining the joint effects of correlation and sample size on D.

Power

Ideally the assessment of the power of a statistic would be based on calculus or at least on the development of tables from many thousands of simulations. Having neither the insight to provide the calculus nor access to the mainframe computer facilities necessary to conduct massive numbers of simulations, a modest number of simulations were conducted on a personal computer. Table 1 provides mean $rde(y)$ and D values at different levels of IV/DV correlation and n. Column headings indicate the mean IV/DV correlations, obtained by the weighting procedure described earlier. The weights ranged from .12 to 6. Row headings reflect varying sample sizes. Altogether there were 144 conditions in this simulation, 18 levels of IV/DV correlations crossed with eight sample sizes. Each condition was replicated 200 times, producing a total of 28,800 model simulations.

Inspection of Table 1 suggests the values of D are reduced as the IV/DV

Table 1

Mean *rde*(y) and *D* Values for 18 Levels of IV/DV Correlation and Eight Sample Sizes

	Average IV/DV Correlation					
	r =.12	r =.24	r =.34	r =.45	r =.52	r =.60
n=25	−08/−02	−15/−09	−22/−16	−29/−23	−36/−31	−41/−37
n=50	−04/−03	−11/−10	−21/−19	−29/−28	−37/−33	−43/−42
n=75	−03/−02	−11/−10	−21/−20	−31/−31	−37/−37	−43/−42
n=100	−04/−03	−12/−10	−21/−20	−31/−31	−38/−37	−43/−43
n=125	−04/−03	−11/−10	−21/−20	−30/−29	−38/−37	−44/−43
n=150	−03/−03	−11/−11	−21/−19	−30/−29	−38/−36	−44/−43
n=175	−04/−03	−11/−11	−20/−19	−31/−30	−39/−38	−45/−45
n=200	−02/−03	−11/−11	−21/−31	−30/−30	−38/−38	−45/−45
	r =.65	r =.71	r =.74	r =.78	r =.81	r =.83
n=25	−43/−42	−45/−42	−44/−39	−42/−38	−41/−36	−39/−33
n=50	−46/−43	−47/−46	−45/−42	−43/−41	−44/−43	−38/−37
n=75	−45/−43	−47/−45	−47/−46	−44/−43	−43/−42	−40/−37
n=100	−47/−47	−49/−48	−47/−45	−46/−46	−44/−41	−40/−39
n=125	−48/−47	−49/−47	−48/−47	−45/−45	−44/−43	−40/−40
n=150	−48/−48	−49/−47	−49/−48	−46/−45	−44/−42	−40/−39
n=175	−48/−47	−49/−50	−49/−49	−46/−46	−43/−43	−40/−40
n=200	−47/−47	−50/−49	−48/−47	−46/−46	−44/−43	−42/−42
	r =.85	r =.87	r =.89	r =.93	r =.96	r =.99
n=25	−36/−31	−33/−28	−26/−22	−23/−18	−15/−10	−10/−04
n=50	−38/−36	−34/−30	−30/−29	−23/−21	−14/−12	−08/−04
n=75	−38/−36	−36/−34	−31/−29	−24/−22	−14/−14	−06/−04
n=100	−39/−37	−36/−35	−30/−29	−24/−23	−14/−13	−05/−04
n=125	−38/−37	−35/−34	−30/−29	−23/−22	−14/−13	−06/−05
n=150	−38/−38	−37/−36	−31/−30	−23/−23	−13/−12	−06/−05
n=175	−39/−38	−35/−34	−32/−30	−23/−23	−13/−13	−06/−05
n=200	−39/−38	−36/−36	−31/−31	−23/−23	−14/−13	−06/−05

Note: Table entries are expressed as *rde*(y)/*D*. Decimal points are omitted to conserve space.

correlation deviates from $r = .7$ and as the sample size decreases. A count of the number of times the *D* value was negative was conducted to provide an estimationn of power. The hit rates were very good. There was near 100 percent accuracy when the IV/DV correlations ranged between $r = .20$ and $r = .90$ and sample sizes were 50 or larger. In most social science applications these restrictions on power would be relatively trivial. There is reason, therefore, to be optimistic about the potential power of corresponding regressions, at least, for the social sciences.

Multiplicative, Subtractive and Reflected Models

The model used thus far has assumed y is generated by adding $x1$ and $x2$. What if nature chooses to create, not by adding, but by multiplying? How will the method of corresponding regressions hold up under such conditions? To answer this question, fifty replications of the basic model were conducted, except that y was calculated by multiplying $x1$ by $x2$. The average correlation between $x1$ and y was $r = .65$. The average value of $rde(y)$ was $-.16$. The average $rde(x)$ was $-.03$. The average D was $-.13$, $sd = .14$, $t = -6.78$, $p < .0001$. The results suggest the method may be less effective in the multiplicative model than in the additive model. Perhaps transformations can be developed to make the method as powerful with multiplicative models as with additive models. Whether or not such adjustments will be forthcoming, the method does appear to work to some extent with multiplicative models.

The usefulness of corresponding regressions with subtractive models may be of particular interest to teleological theorists. Rychlak (1988) has demonstrated the role of dialectical reasoning in decision making, arguing that oppositional thinking can free the person from the constraints of environmental contingencies and intellectual impasses. If people sometimes think by opposing one idea (IV) with another idea (IV) in order to dialectically derive a new synthesis (DV), and if such construction can be reflected in rating scales, it would be possible to trace the history of these thoughts using corresponding regressions. Later in this paper an experiment will be presented that supports the plausibility of such cognitive assessments. For now, however, it is important to discern the usefulness of corresponding regressions with simulated subtractive models.

Recall that the kernel of the method of corresponding regressions was the tendency for the correlations between $x1$ and $x2$ to polarize across the ranges of y. In the additive model the correlation will be positive at the extremes of y and negative in the mid-range of y. With subtractive models, polarization also occurs, but the pattern of correlation is reversed. In fifty replications of a subtractive model, where $y = x1 - x2$, it was found that $x1$ and $x2$ correlated negatively at the extremes of y and positively at the mid-range of y. This should not pose a threat to the validity of the method because the polarization itself, not the valence of the correlations, is the "kernel" of the method. The least squares solution will still be oriented by the extremes. The slope will simply be expressed as a negative value. Consequently, corresponding regressions should still be effective when applied to subtractive models.

To test this effectiveness, fifty replications of a subtractive model, i.e., $y = x1 - x2$, were conducted. The average correlation of $x1$ and y was $r = .71$. The average $rde(y)$ was $-.49$. The average $rde(x)$ was $-.01$. The sd for the difference was .14, $t = -24.60$, $p <. 0001$. Clearly the method was effective when applied to subtractive models.

In the models addressed so far the *x1* and *y* variables have always been positively correlated. This is reasonable in that *y* is determined by *x1*. There are situations, however, when the researcher might find *x1* and *y* negatively correlated. Such negative correlations could easily occur in psychology, since many of the scales used in psychology are arbitrary. How will the method hold up when the values of *y* are reflected, in order to correlate negatively with *x1*? Once again, the "kernel" of the method is not affected. The slope of the regression equation can easily adjust to the arbitrary scale of the *y* variable. Fifty replications of the model $y = 1 - (x1 + x2)$ were generated, however, just to be certain. The average correlation between *x1* and *y* was −.71. The average *rde*(*y*) was −.48. The average *rde*(*x*) was −.01. The average value of *D* was −.47, $sd = .13$, $t = -25.54$, $p <. 0001$. Thus, reflected models present no problem for the method of corresponding regressions.

Uncorrelated Variables and Correlated DVs

A major problem in the development of structural equations is the case of correlated dependent variables. If the method of corresponding regressions is valid, *D* scores should be near zero for variables that are correlated but causally independent. Similarly, low *D* values should also be found for variables that are both uncorrelated and causally independent of one another.

To test the above assumptions fifty replications of two conditions were generated. The first condition simulated correlated dependent variables. In this condition *x* was developed by adding a uniform random variable, *w1*, to another random variable, *z*. *Y* was created by adding a different random variable, *w2*, to the same *z* variable used in the creation of *x*. Thus *x* and *y* had *z* in common and correlate with one another. *Z* is an independent variable for both *x* and *y*. The *x* and *y* variables, however, do not determine one another in any way. In other words, *x* and *y* are correlated dependent variables.

Another condition was developed by simply generating two random variables. The first variable, *x1*, was compared with the second variable, *x2*. This condition simulated uncorrelated, causally unrelated variables.

Fifty replications of the above models were generated. As expected, the *D* values for both conditions did not differ significantly from zero. The mean correlation between *x* and *y* in the correlated dependent variables condition was $r = .50$. The mean *rde*(*y*) was −.12. The average *rde*(*x*) was −.13. The average *D* was .01, $sd = .14$, $t = .56$, n.s. Similarly, the two uncorrelated variables produced results suggesting no dependencies. The average *r* was .01, the average *rde*(*y*) was −.01. The mean *rde*(*x*) was −.01. The average *D* was

0.0. The t-test was of course nonsignificant. Apparently the method is useful in assessing conditions where no causal relations exist, even when the variables in question are correlated.

Recursion and Sequacious Extensions

Simulated Recursive Models

Each of the models has thus far addressed very simple one-way relations. Most of the models that employ structural equation methods, however, contain chains of implication, with one variable causing a second, which in turn causes a third, the third a fourth and so on. In the following simulations, corresponding regressions are examined across varying levels of recursion in order to determine the effectiveness of corresponding regressions across layers of dependencies.

Two conditions were developed to examine levels of recursion. In the first condition D values under three levels of recursion were addressed: $z => x => y$. Three variables were simulated: z, x, and y. The first variable in the chain, z, was a uniform random variable. X was generated by adding another uniform random variable to z. Y was created by adding yet a third random variable to x. Thus x was dependent on z while y was directly dependent on x and indirectly dependent on z. The corresponding regressions were conducted on z and y, with the intention of assessing the power of D when comparing variables at distant ends of a causal chain. There were 100 observations for each of the fifty replications of this model.

A second condition was developed to address the impact of recursion by building a longer chain of causes. In this condition the following chain was simulated: $u => v => w => x => y$. As before, the recursion was generated by adding random variables to sums. U was a random variable, v was the sum of u and a random variable, w was the sum of v and a random variable, x was the sum of w and a random variable, and y the sum of x and a random variable. D was computed from regressions of u and y. As before, there were 100 observations for each of the fifty replications of this model.

Results suggest D was useful in both conditions of recursion. The mean correlation between z and y in the first condition was $r = .60$. The mean $rde(y)$ was $-.30$. $Rde(x)$ was, on average, .01. The average D in the first condition was $-.31$, $sd = .13$, $t = -16.70$, $p < .0001$. The mean correlation between u and y in the second condition was $r = .43$. The mean $rde(y)$ was $-.17$. The mean $rde(x)$ was $-.02$ and the mean D was $-.15$, $sd = .13$, $t = -8.31$, $p < .0001$. Results suggest the method of corresponding regressions is useful in detecting asymmetrical relations across several levels of recursion.

Sequacious Extentions

It is important to consider the meaning of recursion from a formal cause perspective. It would be easy to mistakenly associate chains of causation with temporal progressions. Indeed, the search for temporal sequences is a major goal in traditional path analysis, since temporal precedence is the most trusted tool in efficient cause inference. The consequent billiard ball model of causation is very appealing in its simplicity. Formal cause recursions, however, are in a sense more fundamental than temporal sequences. When $x1$ and $x2$ are added, they simply equal y. This generation is atemporal, just as $2 = 1 + 1$, is independent of time. Rychlak (1981) explains such atemporal sequences of dependency by reference to logical terms. "A precedent meaning is one that goes before others in logical arrangement without regard for passage of time. It is not an 'antecedent' occurring 'first' over the flow of time in relation to a 'consequent.' A sequacious meaning is one that necessarily (logical extension) brings forward the meaning of precedents. There is a 'slavish compliance' of sequacious meanings on their precedents, so that they always follow through or elaborate on what has been framed initially. The order here is one of formal causation, obtaining as a pattern outside of time's passage, and not as a consequent being impelled in an efficient-cause sense by an antecedent over the flow of time" (p. 427). Rychlak's argument does not, of course, suggest that there are no logical patterns across time. Formal cause patterns can sometimes be useful in delineating temporal sequences. Researchers should, however, be careful to avoid equating logical sequences with historical sequences. As long as temporal sequences conserve tautologically earlier links in the chain, the structure of a known temporal sequence may be delineated. If, however, a logical break occurs in the chain, the formal cause sequence can no longer reflect the historical sequence. For example, people do not possess all of their ancestor's genes. Because of the limited transfer of chromosomes at conception, a pattern of genetic disinheritance emerges across generations. After so many recursions there is the possibility that no commonality exists between the temporal precedents and consequents. The "sins of the fathers" cast their shadows across a limited number of generations. Nevertheless, for a while, the formal cause patterns do reflect the history of a family of variables. This was supported by the effectiveness of the corresponding regressions when applied to a finite number of simulated levels of recursion. The fact that the D values reduced with larger numbers of recursions suggests, however, that a fading or disinheritance factor must be taken into account when applying corresponding regressions.

A second type of logical break in a sequacious extension should be addressed because of its relevance to Aristotle's notion of final cause. The genetic disinheritance described above need not be a passive process, driven primarily by the limited capacity of a system to incorporate information. A person could actually choose to break a pattern. The logic of a sequacious extension could be purposely altered in order to transcend or disinherit certain precedents. This type of break would provide a degree of freedom to the person's construction processes and could thus have profound implications for the teleological theorist. By using something like path analysis in combination with corresponding regressions, a researcher could conceivably locate specific points where a precedent drops out, abruptly, from a sequacious extension. This could be useful in understanding processes involved in therapy, problem solving, and creativity. This theme will be taken up again in a later section. For now, the investigation turns to more concrete demonstrations of the validity of the method of corresponding regressions.

Applications of the Method to Real Data

Examples from the Natural Sciences

Thus far the data have been simulated. The use of simulated data has provided a level of clarity and control that could not have been obtained from real data. To be generally useful, however, the method of corresponding regressions must be applicable to real data. The difficulty in using real data to validate the method of corresponding regressions is the very reason for the need for the method. The assessment of causal relations among psychological, sociological, economic, and similar phenomena can be very difficult and the final word will definitely not be said in this paper. In a sense this study has been an attempt to validate an objective method by a subjective method – that is, by intuition. The following examples illustrate the breadth of potential applications of the method. They also illustrate the author's sometimes vague and even arbitrary causal theorization. Therefore, the examples are not presented as demonstrations of the validity of the method, rather, as enticements to those who would further investigate the validity of the method.

An example from physical science will be used to begin the exploration of corresponding regressions in real data. The experiment consisted of pouring different amounts of water from two test tubes into a beaker, in order to cause the level of water in the beaker to change. The independent variables were the amounts of water poured from each test tube. The dependent variable was the amount of water in the beaker, collected from both test tubes. Thus the amounts of water in the test tubes directly caused the amount of water in the beaker. The amounts of water in the test tubes were counterbalanced so that ten levels in tube $x1$ were combined once with each of ten levels in tube $x2$. There were thus 10×10 or 100 observations of the variables. The analysis of corresponding regressions was performed on the water levels in test tube $x1$ and beaker y.

To apply the method of corresponding regressions to real data it is convenient to develop a simple metaphor. Imagine there is a box with two doors, *a* and *b*. Observations of two variables are introduced to one door or the other. The goal is to develop a procedure that indicates the causal relation between the values passed to the doors without knowing which door received the dependent or independent variable. This can be done simply by letting door *a* receive the first predictor variable while door *b* receives the second predictor variable. As in the simulations, *rde* correlations for the first predictor are then subtracted from those for the second predictor. If the subsequent value of *D* is significantly negative then the variable passed to door *a* was an independent variable and the variable passed to door *b* was a dependent variable. A positive value of *D* would suggest the variable passed to door *b* was the IV. If *D* is not significantly different from zero, the variables passed to doors *a* and *b* have no causal relationship. For clarity's sake, the hypothesized independent variable will always be passed to door *a*. The expected *D* value will thus be a negative value.

In the experiment consisting of water levels in test tubes and a beaker, levels of tube *x1* were passed to door *a* – that is, tube *x1* water levels served as the predicted independent variable. Water levels in the beaker were passed to door *b*, thus serving as the predicted dependent variable. The correlation between water levels in tube *x1* and the beaker was $r = .71$. The value of *rde* for door *b*, *rde*(*b*), was −.52. *Rde* for door *a*, *rde*(*a*), was 0.00. The value of *D* was thus −.52. Consulting Table 1 suggests this value of *D* is highly significant. The method of corresponding regressions thus properly assessed the cause and effect variables in this simple applied example.

A somewhat more complicated example of causation is found in a study concerning electronics. Sidik, Leibecki, and Bozek (1980) collected data on the relationship between temperature and cycles to failure in silver-zinc batteries. Temperature was assumed to decrease the number of cycles to failure. The original data was published in Johnson and Wichern's (1982, p. 356) multivariate statistics text. Temperatures and cycles to failure of twenty-four experiments were entered into the metaphorical "box," assuming temperature was the IV. The correlation between cycles to failure and temperature was $r = .66$. The value of *rde*(*b*) was −.15. The value of *rde*(*a*) was .20. The value of *D*, −.36, was consistent with the presumed causal relationship. Similar data were found in Kleinbaum and Kupper's (1978, p. 156) text on multivariate statistics. Yoshida (1961) examined oxygen consumption rates of 47 groups of wireworm larvae at different temperatures. Again, the experimenters assumed temperature was the IV. Oxygen consumption rate was assumed to be the DV. The data were passed to the metaphorical box. The correlation between oxygen and consumption and temperature was $r = .90$. Again, the value of *D*, −.37, with *rde*(*b*) = −.38 and *rde*(*a*) = −.01, suggested agreement between conventional scientific wisdom and the method of corresponding regressions.

Examples from the Social Sciences

The first data in the social science area investigated by the writer came from the *Statistical Abstract of the United States: 1988*. Federal expenditures on human resources and national defense for the thirty years sampled since 1950 were examined. The writer presumed defense expenditures determined human resource expenditures. The writer's lack of knowledge in economics, however, made this prediction little more than a guess. In fact, the results suggested the opposite causal pattern. The correlation between the expenditures was −.79. The *rde*(*b*) was .23. The *rde*(*a*) was −.30. The value of *D* was thus .53. The method thus suggests the national defense expenditures were dependent on human resource expenditures. The meaningfulness of these findings must be judged by those with a good understanding of economics.

A second data set was also taken from the *Statistical Abstract* for 1988. The relationship between crime rates per 100,000 and expenditures per capita on criminal justice activities was examined for the 50 states and the District of Columbia. Crime rates were assumed to be the dependent variable. The statistics were inconclusive. The *rde*(*b*) was .04. *Rde*(*a*) was .36. *D* was −.32. The correlation between expenditures and crime rate was .53. Notice, however, that both of the *rde* correlations are positive. The *D* in this experiment is unlike that of any other encountered in the previous investigations. Since the method of corresponding regressions is based on negative *rde* correlations, and since the *D* in this example is based on positive correlations, the best conclusion, assuming the method is valid, is that there is no causal relationship.

A third venture into the social science realm proved to be equally perplexing. In this study data provided by Johnson and Wichern (1982, p. 305) were analyzed. The data concerned assessed values and selling prices of houses. The assessed values were taken as the IV. Actual selling prices were assumed to be the dependent variable. Results contradicted these assumptions. The correlation between the two variables was $r = .85$. The values of *rde*(*b*), *rde*(*a*), and *D* were respectively: .33, −.35 and .68. The data suggest selling prices determined assessed values.

A fourth social science application of the method addressed an issue in personality psychology. Personal construct theory (Kelly, 1955) suggests that personality is shaped by the constructs a person uses to anticipate events. Most construct theorists assume constructions of self, family, and friends are of particular relevance to the person's "core" functioning. In a re-analysis of data reported in detail by Chambers and Epting (1985), the relationship between construct logical consistency and neuroticism scores on Eysenck's E.P.I. was addressed. Neuroticism was assumed to be causally dependent on perceptions of self, family, and friends. There were 40 subjects in the experiment, all of them college students. Results indicate that logical consistency is dependent on neuroticism, $rde(b) = .01$, $rde(a) = -.20$, $r = .49$. The data suggest the inconsistencies were probably symptomatic of some more underlying instability, rather than being the source of the misery reflected in the E.P.I.

The Form of a Final Cause

Earlier it was pointed out that Rychlak (1981) has drawn attention not only to formal causes but also to final causes. Final causes concern telic sources of change. A final cause is the development of a sequacious extension for the purpose of obtaining some outcome. The final cause is not the sequacious extension itself. The sequacious extension is merely the structure of a set of choices. Indeed, as was discussed in the earlier section concerning recursion and sequacious extensions, one could choose to break the pattern of a sequacious extension to serve some purpose. It is reasonable to assume, however, that much can be learned about the processes of decision-making by studying sequacious extensions in cognition, whether these extensions are broken or not. In fact, Rychlak (1988) made final causes and formal causes central to his theory of logical learning. This theory suggests that an understanding of mind and behavior must proceed from assessments of the intentions of the individual. Learning and decision-making are viewed as the willful development of sequacious extensions of intentions, with affective evaluations frequently serving as the premises that are sequentially elaborated through the individual's purposeful behavior. The goal of the final study will be to explore the form of such telic sequences, attempting to delineate the chain of thoughts making up intentional behavior. Although no substantive contributions to an understanding of decision-making are expected from the present study, the application of corresponding regressions to sequences of thoughts may prove suggestive for future research.

The basic causal model employed in most of the simulations in this paper has been a linear combination of two variables, with the consequent y values depending on the precedent x values. If people employ similar mental arithmetic, the method of corresponding regressions should reflect the sequence of dependencies in thoughts. To test this notion, 32 undergraduate volunteers were asked to list 16 activities that they intended to pursue in the coming summer. The activities were elicited in a manner such that each activity expressed some unique combination of desire and opportunity. Four levels of desire and four levels of opportunity were crossed, in a fashion reminiscent of the procedure used in the test tubes and beaker experiment described earlier. To do this the subjects first were asked to list an intended activity that they had a high (4) desire to pursue and a high (4) expectation of finding opportunities to pursue. Next, subjects listed a highly desirable (4) activity for which there was a moderate (3) opportunity of involvement. Subjects went on to list a highly desirable (4) but moderately inconvenient activity (2) and a highly desirable (4) but highly inconvenient (1) activity. Then the process continued across levels of opportunity for moderate (3) levels of desire, as well as across moderate (2) and high (1) levels of non-desire. Thus uniformly distributed levels of desire and opportunity were elicited. Subjects were asked to list only activities that they could freely choose to pursue. Since many of them felt compelled to work over the summer, they were encouraged not to use this activity. Subjects were further asked to study their attributions of desire and opportunity and on the basis of these judgements, to indicate their expected levels of involvement in the activities. The subjects were asked to find four activities in which they intended to be most involved and to assign these activities the rating (4). The four activities they were next most likely to pursue were assigned the rating (3). The next four were given the rating (2). Finally, the four activities in which the subjects anticipated the least involvement received the rating (1).

The goal of the above research design was to elicit uniformly distributed ratings of desire, opportunity and involvement across a list of intended activities. The subjects were encouraged to develop sequacious extensions by taking into account their judgements of desire and opportunity in predicting involvement. They were specifically instructed to make their predictions of involvement a function of the combination of desire and opportunity, thus developing the model: involvement = desire + opportunity. Most subjects reported after the experiment that they had taken levels of desire and opportunity into account in estimating involvement. There is, therefore, a strong likelihood that the design successfully elicited sequacious extensions from the subjects, with desire and opportunity being extended through the judgements of involvement. If, in fact, involvement was dependent on desire and opportunity in the subjects' thoughts, the method of corresponding regressions should demonstrate this dependency. Given the emphasis Rychlak (1988) places on affective assessment in logical learning theory, the ratings of desire were used as the IV and passed to door *a* of "the box."

The mean correlation between desire and involvement across the 16 activities for the 32 subjects was $r = .37$. The average D value was $-.05$, $sd = .13$, $t = -2.12$, $p < .05$. Apparently the subjects did generally form their estimates of involvement by combining the levels of desire and opportunity, although the average correlation between desire and involvement at .37 suggests the combination was not a simple linear function. Perhaps the failure of this correlation to obtain the .70 level was a function of broken sequacious extentions. Perhaps this break was distinct and due to willful choices to abandon the additive model. Perhaps these breaks were gradual or fading and reflect the kind of information "disinheritance" to be expected in simple forgetting. Only further research will tell. The average size of *D* presents some concern as well. The average size of the IV/DV correlation falling lower than .7 and the relatively small sample sizes, i.e., 16, would tend to lower the power of *D*. Nevertheless, the two-tailed probability criterion was exceeded at a modest level of significance, suggesting the method of corresponding regressions reflected sequacious dependencies in decision-making. Thus, the method appears to have disclosed the form of a final cause.

The results of the final experiment are intriguing because they suggest that the method might be useful in more substantive explorations of the mind and behavior. One might ask 100 subjects to rate their mothers and wives on 50 adjectives. Would the subjects' views of their wives be dependent on the views of their mothers, especially the perceptions of those subjects who demonstrate signs of Oedipus complex? What would Freud have given to be able to conduct such an experiment? Would Adorno have been interested in determining whether or not authoritarian Americans' ratings of the US foreign policy were dependent on their ratings of the Sermon on the Mount? What could the notions of sequacious extention and disinheritance contribute to memory research or to an understanding of repression? Could the method provide assessments of personal construct hierarchies that are inaccessible through factor analysis? How would Plato, Aristotle, Leibniz or Kant have made use of such a method? Is it possible that philosophical and religious questions concerning the structure of the mind, ethics and religious experience could be answered with the method? How often have philosophers, historians, religious leaders, and psychologists claimed that the actions of a group of people were dependent on their views of this or that? Of course it is too early to claim to have stumbled upon a psychometric Holy Grail. The sample and the effect were modest. Still, the final cause experiment did represent a sequacious extention of the rationale of the earlier studies. With further research, the teleologist may finally shake the constraints of efficient cause designs. The billiard ball model may give way, perhaps, to a cathedral model, with the relations and dependencies between nave and transepts, apse and chevet, being scientifically demonstrable. Such models of the psyche and society have been contemplated for centuries. Perhaps the opportunity has arrived.

References

Chambers, W.V. (1968). Inferring causality from corresponding variances. *Perceptual and Motor Skills, 63*, 475–478.

Chambers, W.V., and Epting, F.R. (1985). Personality and logical consistency of personal constructs. *Psychological Reports, 57*, 1120.

Cook, T.D., and Campbell, D.T. (1979). *Quasi-experimentation*. Boston: Houghton Mifflin.

James, L.R., Muliak, S.A., and Brett, J. (1982). *Causal analysis: Assumptions, models and data*. Beverly Hills: Sage Publications.

Johnson, R.A., and Wichern, D.W. (1982). *Applied multivariate statistical analysis*. Englewood Cliffs, New Jersey: Prentice-Hall.

Kelly, G.A. (1955). *The psychology of personal constructs*. New York: Norton.

Kleinbaum, D.G., and Kupper, L.L. (1978). *Applied regression analysis and other multivariate methods*. North Scituate, Massachusetts: Duxbury Press.

Rychlak, J.F. (1981). *A philosophy of science for personality theory*. Malabar, Florida: Krieger Publishing Company.

Rychlak, J.F. (1988). *The psychology of rigorous humanism*. New York: New York University Press.

Sidik, S., Leibecki, H., and Bozek, J. (1980). Failure of silver-zinc cells with competing failure modes – Preliminary data analysis. *NASA Tehcnical Memorandum 81556*. Cleveland, Ohio: Lewis Research Center.

Simon, H.A. (1977). *Models of discovery*. Dordrecht, Holland: R. Reidel.

Statistical Abstract of United States. (1988). Washington, D.C.: U.S. Bureau of the Census.

Statistical Yearbook. (1982). New York: Publishing Division, United Nations.

Yoshida, M. (1961). *Ecological and physiological researches on the wireworm, Melanotus caudex Lewis*. Tokyo: Iwata Shizouda Pref.