Assumptions and Interventions of Probabilistic Causal Models

Chong Ho Yu, Ph.D.

October, 2002

RUNNING HEAD: Causal assumptions

Paper presented at Arizona Educational Research Organization
15th Annual Conference

Correspondence:

Chong Ho Yu, Ph.D.

Psychometrician

PO Box 612

Tempe AZ 85280

USA

Email: asumain@yahoo.com.hk

Website: http://seamonkey.ed.asu.edu/~alex/

Abstract

Causality is an intriguing but controversial topic in philosophy, statistics, as well as educational and psychological research. By supporting Causal Markov Condition and Faithfulness Condition, Clark Glymour attempted to draw causal inferences from structural equation modeling. According to Glymour, in order to make causal interpretation of non-experimental data, the researcher must have some type of manipulation, rather than conditioning, of variables. The Causal Markov Condition and its sister, the common cause principle, provide the assumptions to structure relationships among variables in the path model and to load different variables into common latent constructs in the factor model. In addition, the Faithfulness Condition rules out those models in which statistical independence relations follow as a result of special coincidences among the parameter values. The arguments against these assumptions by Nancy Cartwright as well as those for these assumptions by James Woodward will be evaluated in this paper.

Assumptions and Interventions of Probabilistic Causal Models

Chong Ho Yu, Ph.D.

Introduction

Causality is an intriguing but controversial topic in philosophy, statistics, and the social sciences. Since the introduction of Pearson' s Product Moment Correlation Coefficient, many statisticians and social scientists have been conducting research based upon association. For a long time the question about whether quantitative methodologies could lead us to causal inferences has remained unsettled.

There are some sound reasons to justify why people are skeptical toward causal inferences yielded by statistical models. Yule (1926) pointed out that sometimes we could get nonsense-correlations between time-series. For instance, if you plot GNP, educational level, or anything against time, you may see some significant correlation. On the other hand, even though bad research studies exist, it does not mean that we should abandon the endeavors altogether. In recent years, both Glymour and his CMU group (Glymour, 1982, 1983; Glymour, Scheines, Spirtes, & Kelly, 1987; Glymour, 1999; Glymour & Cooper, 1999) have been devoting efforts to the TETRAD project in an attempt to affirm causal inferences based upon correlational information and non-experimental data. Not surprisingly, many scholars have voiced either their support or objections to Glymour et al.' s approach.

Interestingly enough, numbers per se could not determine whether causal information could be extracted from the data or the mathematical model. Basically, both proponents and opponents of using statistical approach in causality utilize the same numeric information. For instance, structural equation modeling, the causal model endorsed by Glymour and Pearl (2000), is composed of a measurement model and a path model. In a measurement model, Pearson' s Product Moment Correlation Coefficient, which is assocational in essence and a-causal in origin, is used for factor analysis. In addition, today the widely used hypothesis testing by statisticians and social scientists is a fusion of Fisher, Pearson, and Neyman' s models. As mentioned before, Pearson accepted association only and de-emphasized causality. Regardless of whether you believe in causality or not, you may still conduct hypothesis testing, run Pearson' s Correlation Coefficient, and/or do factor analysis, unless you totally reject quantitative methods.

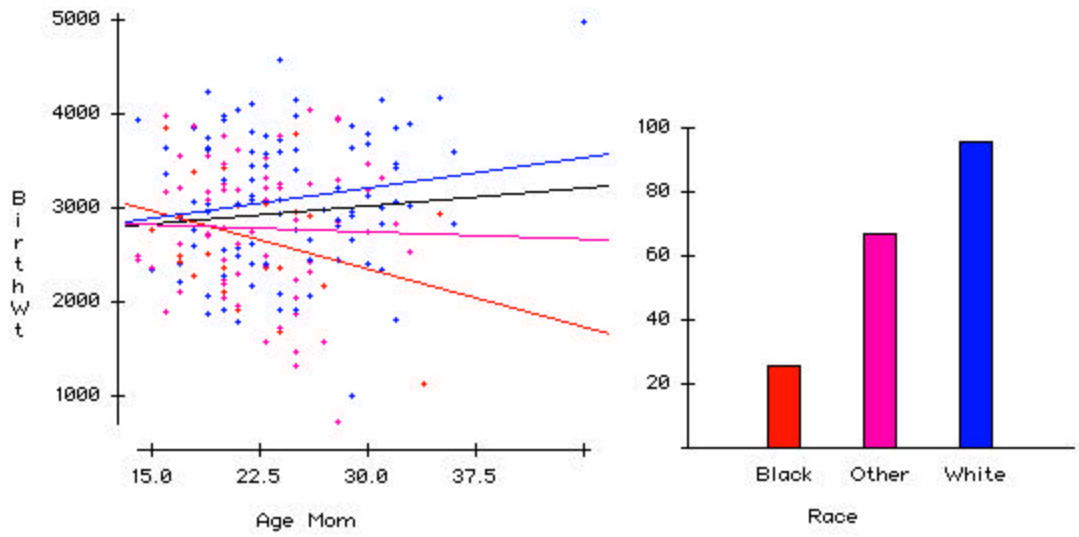
If numbers and mathematics alone could not settle the debate of causality, then where could

we go to investigate the problem? I believe that the problem is concerned with the philosophical aspects, such as the unproved assumptions of statistical modeling. In this paper, two major assumptions of Glymour' s TETRAD will be discussed. The arguments against these assumptions by Nancy Cartwright as well as those for these assumptions by James Woodward will be evaluated.

Conditioning

As mentioned in the beginning, when researchers compute the association of observational (non-experimental) data, sometime the relationships might seem to be nonsense. In order to gain more insights, careful statisticians might partition the data by grouping variables or other lurking variables. This kind of activity can be considered "conditioning." For example, in a research study regarding the relationship between the birth weight of babies and the age of mothers (an example dataset included in DataDesk, Data Description, Inc., 1999), the regression slope using the full dataset (see the black line and the blue bar in Figure 1) indiciates that as the age of mothers increases, the birth weight of babies increases. This relationship is counter-intutitive because usually as the mother gets older and older, the chance to give birth to a healthy baby is lower and lower.

Figure 1. Relationships among birth weight, age of mother, and race



However, when the dataset is partitioned by a grouping variable, race, the issue becomes more complicated. The positive relationship between birth weight and age is true among whites. For

blacks, the relationship is negative (see the red line and the red bar in Figure 1), while for other ethnic groups the regression slope is almost flat, and therefore no significant relationship is implied. Please keep in mind that this study is non-experimental for the researcher did not manipulate age, race, and birth. Conditioning this kind of observational data always faces this problem: No broad generalization about relationships could be firmly made because further conditioning and partitioning may reverse the relationship discovered in the aggregate dataset.

Intervention and Manipulation

According to Meek and Glymour (1994), computing probabilities by conditioning on an event is very different from computing probabilities upon an intervention to bring about that event. While talking about intervention, readers may get an impression that Glymour was talking about conducting experiments, in which human interventions are imposed on various scenarios. Indeed, in Glymour's view, intervention does not necessarily happen at the data collection stage. At the data analysis stage, data manipulation and model building can also be viewed as a different kind of intervention.

Meek and Glymour compared the Fisherian tradition with their own work to show the continuity between both. Fisher's design of experiment could achieve two objectives: (1) To ensure that treatment assignment has no common causes and are independent if treatment has no effect on outcome; (2) to determine a definite joint probability distribution for treatment and outcome under the assumption of no effect (null hypothesis). On one hand, Fisher's design of experiment requires randomization of group assignment to rule out common causes. On the other hand, Meek and Glymour asserted that causal claims entail claims about intervention or manipulation. If the research study is not experimental, then how could the logic of the Fisherian school be applied to causal inferences of non-experimental data? Spirtes, Glymour & Scheines (1993) proposed that two assumptions could be employed to bridge the gap between the causal structure and the non-experimental data: the Causal Markov Condition (CMC) and the Faithfulness Condition (FC). In their view, equipped with these two assumptions, researchers could draw causal inferences as if intervention or manipulation had been made to the data.

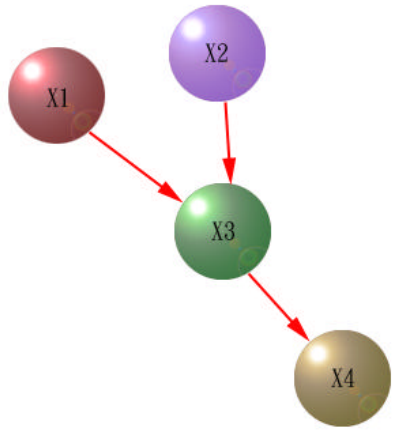
Causal Markov Condition

In a causal model, joint probability distribution over the variables must satisfy CMC (Druzdzel & Glymour, 1995). In CMC, each variable is probabilistically independent from its non-descendants, conditional on its parents. In Figure 2, suppose that X_1 and X_2 are

probabilistically independent from each other, and they both contribute to the effect of X_3. X_4 is independent from X_1 and X_2, conditional on X_3. If X_1 and X_2 were not probabilistically independent, the model would be problematic. For example, in a regression model when independent variables are highly correlated, the problem of multicollinearity exists and the model is not interpretable. The Causal Markov Condition is the assumption of the path model, in which relationships among variables are structured. The path model is one of the components of the structural equation model adopted by causationists.

The Causal Markov Condition also implies the common cause principle proposed by Reichenbach (1956) and advocated by Glymour and his colleagues (Glymour, 1982; Glymour, Scheines, Spirtes, & Kelly, 1987). According to the common cause principle, if a system of variables satisfies the Markov Condition, and they have a high degree of association, then there exists a latent construct (factor) causing them. The common cause principle is the underlying assumption of the factor model, which is also a building block of the structural equation model.

Figure 2. Example of the Causal Markov Condition

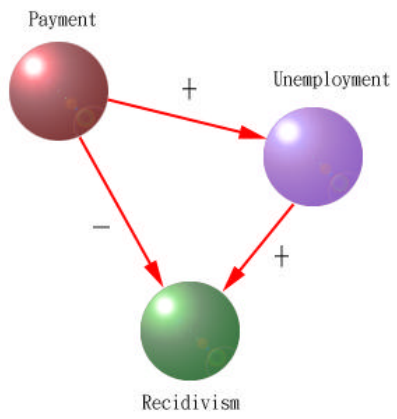


Faithfulness Condition

According to the faithfulness condition, statistical constraints arise from structure, not coincidence. As the name implies, FC supposes that probabilistic dependencies will faithfully reveal causal connections. In other words, all independence and conditional independence relations among observed variables are consequences of the CMC applied to the true causal structure. For example, a research study (cited in Glymour, 1987) indicates that providing

financial aid to released prisoners did not reduce recidivism. An alternate explanation is that free money discourages employment, and unemployment has a positive effect on recidivism while financial aid tends to lower recidivism. As a result, these two effects cancel out each other (Figure 3). However, the faithfulness condition rules out this explanation.

Figure 3. Example of the Faithfulness Condition.



Manipulation Theorem

Meek and Glymour (1994) proposed that when probabilities satisfy CMC and FC, and when the intervention is ideal in the sense of manipulation, casual inferences are legitimate. This notion is termed the "manipulation theorem." To be specific, given an external intervention on a variable A in a causal model, the researcher can derive the posterior probability distribution over the entire model by simply modifying the conditional probability distribution of A. If this intervention is strong enough to set A to a specific value, the researcher can view the intervention as the only cause of A. Nothing else in the model needs to be modified, as the causal structure of the system remains unchanged.

To implement this theorem, Glymour and his CMU group developed a software-plugin named TETRAD to manipulate/intervene on structural equation models by searching all possible paths among variables (manipulation by "what-if"). It is important to note that TETRAD is not something entirely new. Popular structural equation modeling software applications such as

LISREL and EQS have their own automatic path searching algorithms. Nevertheless, Ting (1998) found that the hit rates (the success rates of uncovering the right causal structure) of TETRAD' s automatic search procedure reach 95% for large samples (n=2000) and 52% for small samples (n=200), which are far higher than those offered by LISREL and EQS.

Cartwright' s Arguments against Glymour' s Ideas

Empiricist view: No causes in, no causes out

Many philosophers are opposed to the preceding idea. Due to space constraints, this paper will concentrate on Nancy Cartwight only. Cartwright (1999) emphasized the point of "no causes in, no causes out." (p. 39) To be specific, there is no way to get casual information from equations and associations. New causal knowledge must be built only from old, empirical causal knowledge. In other words, the empiricist' s rule embraced by Cartwright is that the relevant data are the data that will fix the truth or falsity of the hypothesis, given the other known facts. Glymour et al. included all possible combinations of variables and paths in the model and then irrelevant ones were eliminated. Cartwright questioned that if relevant variables and genuine causes are not included at the beginning, then this elimination approach is useless. For these reasons, Cartwright strongly criticized Glymour et al.' s theory:

> "Because Glymour, Scheines, Kelly, and Spirtes employ the hypothetico-deductive method, they must proceed in the opposite order. Their basic strategy for judging among models is two-staged: first list all the relevant relations that hold in data, then scan the structures to see which accounts for the greatest number of these relations in the simplest way. That means that they need to find some specific set of relations that will be relevant for every model. But, from the empiricist point of view, no such thing exists." (p.78)

In questioning the applicability of CMC, Cartwright (1999) used a classical example to argue that researchers may take the risk of confusing a co-symptom with a cause: In R.A. Fisher' s opinion, smoking does not cause lung cancer. Rather, smoking and lung cancer are caused by a common cause: a special gene that increases the tendency to smoke and to get cancer. Not surprisingly, Cartwright asserted that to investigate a hypothesis like this, one must conduct a randomized experiment instead of counting on CMC and mathematical intervention of non-experimental data.

Actually, Glymour and his CMU group do not rely on equations alone. Rather, they still use empirical data though the data are not non-experimental. It seems that in Cartwright' s view,

non-experimental data are not "empirical" enough. First, it is a well-known fact that most data in astronomy and geology are not experimental, yet many conclusions in these disciplines are qualified to be causal inferences. Second, according to the abductive logic, new knowledge does not necessarily arise from old, empirical knowledge (Yu, Behrens, & Ohlund, under review). Nonetheless, these popular arguments will not be repeated here. Instead, the discussion will be focus on the nature of empirical data.

A simple definition of empirical data is data that are collected through sensory input and could be verified by sensory channels or logical means. When observational data of various variables are measured and computed, are their statistical properties empirical? Assuming that the data values of these variables indicate a high degree of internal consistency and a single dimension, and thus these variables satisfy the common cause principle and are collapsed into a single factor, can we regard properties such as "internal consistency" (in psychometric sense, not in logical sense) and "unidimensionaity" empirical? My answer is "yes" because they are absolutely verifiable.

Further, assume that I took an IQ test and achieved a score of 200; is the psychometric attribute "high intelligence" empirical? According to strict empiricists, the answer is "no" because the score is not obtained by repeated experiments. Gaining a high score in one single test could be due to pure luck. Right before the test is administered to me, I might take Ginkgo Biloba or read a book carrying IQ test items that are similar to the test. To estimate my IQ score in a scientific manner, I have to retake the same test several times and to demonstrate a high degree of stability of test score over time. However, in many experimental studies subjects are tested or measured just once. In theory, the subjects' memory about the test should be wiped out so that no carry over effect is present when subjects are retested. Needless to say, it is impossible and unethical to erase people's memory. Indeed, reliability of many experimental scores is established by mathematical modeling. To be specific, by thought experiment the true score model assumes that if the same person takes the same test over and over, error scores would scatter around the true score, and the observed score is the composite of the true score and the error score. Hence, mathematical models are applied to minimize the error score (Yu, 2001). Please keep in mind that "manipulation" of the test score is carried out during the data analysis.

Last but not least, it is doubtful whether objecting that a model may leave out some genuine causes or relevant variables and so rejecting the method could help scientific progress at all. First,

who could affirm that all relevant variables are included in the model except the omnipotent God? Second, is it really necessarily to include all relevant variables? In defense of his standpoint, Glymour (1999) wrote,

> "Cartwright is perhaps correct that the whole truth about anything is very complex; but, quite properly, science is seldom interested in the whole truth, or aided by insistence upon it. In my view, an inquiry that correctly found the causes of most of the variations in a social phenomenon and neglected small causes would be a triumph." (p.59)

Causal Markov Condition, probabilistic causation, and Simpson's paradox

In addition, while Glymour et al. based their causal modeling on probability, Cartwright (1999) believed that causal laws cannot be reduced to probabilistic laws and thus CMC is questionable. According to Cartwright, "probabilities may be a guide to causes, but they are like a symptom of a disease: there is no general formula to get from symptom to disease" (p.243). Nevertheless, she did not reject CMC altogether. Rather she pointed out that there is not a universal condition that can be imposed on all causal structures. By citing the Simpson's Paradox (1951), in which the conclusion drawn from the aggregate data is contradicted by the conclusion drawn from the contingency table based upon the same data, Cartwight (1983, 1999b) asserted that universal causal inferences are misleading. The so-called causal relationship is always confined to a particular population. For instance, in England once a 20-year follow-up study was conducted to examine the survival rate and death rate of smokers and non-smokers. The result implied a significant positive effect of smoking because only 24% of smokers died compared to 31% of non-smokers. However, when the data were broken down by age group in a contingency table, it was found that there were more older people in the non-smoker group (Appleton & French, 1996). Based on the Simpson's Paradox, Dupre and Cartwright (1988) suggested that there are only probabilistic capacities, but no probabilistic causal laws at all. In Cartwright's view, causal explanation depends on the stability of capacities. In contrast to probabilistic causation that is relative to grouping variables, capacities remain the same when removed from the context in which they are measured.

Inconsistent results happen all the time. If we reject probabilistic causation because there is inconsistency, many research projects would become impossible. As a matter of fact, the discovery of Simpson's paradox does not discourage researchers from drawing generalizations. Instead, different techniques have been employed by statisticians and social scientists to counteract the

potential threat of Simpson's Paradox. For example, by simulation, Hsu (1989) found that when the sample size is small, randomization tends to make groups become non-equivalent and increase the possibility of Simpson's Paradox. Thus, after randomization with a small sample size, researchers are encouraged to check the group characteristics on different dimensions (e.g. race, sex, age, ...etc.), and re-assignment of group membership is recommended if non-equivalent groups exist. Further, to avoid the Simpson's Paradox, Olkin (2000) recommended that researchers employ meta-analysis rather than pooling. In pooling, data sets are first combined and then groups are compared. As a result, the conclusion drawn from the combined data set could be misleading, while insights about the research question are hidden in partitioned data. On the contrary, in meta-analysis (Glass, 1976; Glass & Smith, 1981; Hunter & Schmidt, 1990) groups in different data sets are compared first in terms of effect size, and then the comparisons are combined to infer a generalization (Table 1). In other words, the information of partitioned datasets is given consideration first.

Table 1. Example of meta-analysis

Study ID	Experimental group mean	Control group mean	Effect size	Correlation coefficient
1	109	100	1.2	0.90
2	215	200	0.6	0.80
3	309	300	0.9	0.80
Average			0.9	0.86

Counter-example of Faithfulness Condition: No models in, no causes out

Cartwright (1999a) was skeptical toward the universality of FC. In FC, it is not acceptable to have two equally powerful causal effects cancel out each other. Nonetheless, Cartwright gave one counter-example: Consider the case of fiber optics. Low-loss optical fibers can carry information at gigabits per second over a long distance. But pulse broadening effects inherent to fiber optics can also smear data as they travel along the cable. In this case, the same original source emits two opposite effects that tend to cancel out each other. In order to make fiber optics useful,

network engineers must apply complicated engineering schemes to enhance the first effect and suppress the latter simultaneously. Thus, Cartwright (2001) argued that the researcher must gather the background information about the causal structure under study instead of blindly following FC under all circumstances. Any conclusions the researcher draws about causal inferences based upon FC can only be as secure as our models of that structure and its operation. That's why Cartwright insisted on "no models in, no causes out" (1999, p.17).

Actually these so-called "counter-examples" can be found everywhere. Take another networking scheme as an example: Consider Ethernet and Category 5 cabling. Ethernet follows a bus topology, and thus any one in the network could send out any data packet at any time. As a result, data packets may collide with each other. Again, engineers must apply complicated methods to make Ethernet useable. Medicine is another example. Last year I was very sick for a long time in spite of visiting clinics over and over. Later I was told that my illness was prolonged because while the medicine that I took cured a certain kind of disease, it also weakened my body and made me vulnerable to other diseases.

Nevertheless, when we examine these cases carefully, we could find that there were no violations of FC. In the networking examples, the positive and negative effects must be exactly equal in intensity so that when networking cables are installed, nothing would happen. But it is not true. In the first case, the signals are indeed sent into the fiber optics but they disintegrate as they travel a longer distance. In the second case, the Ethernet network is still functioning except that data collisions happen all the time. In the example of taking medicine, "feeling sick," like "economic problem" and "social ill," is too ambiguous to be an effect. Indeed feeling sick due to one symptom is different from feeling sick due to another. In other words, the positive and negative effects do not cancel out each other. If the government increases spending to increase the employment rate, the inflation rate may increase, too. But it is incorrect to say that the two effects of government spending cancel out each other because "economic problems" still exist.

Further, without the faithfulness assumption, any model could always be defensible by the argument of "canceling-out." When the welfare program administered to released prisoners is not successful in reducing recidivism, it is said that both the encouraging effect and the discouraging are at work. The same approach can be used to explain any failure or ineffectiveness. If a doctor prescribes the wrong medicine to me and thus my illness is never cured, he could also argue that his medicine works but the drug has a side effect to make me feel bad. A good treatment should be

a robust one. By applying FC, researchers are forced to give a verifiable and clear-cut causal conclusion instead of explaining away failure.

Woodward's Arguments for Glymour's Ideas

Interventions

James Woodward proposed many interesting ideas that are supportive to Glymour's theory. Two important aspects are his view to interventions and invariance. As most people notice, Glymour's idea of intervention is not the same as the conventional sense. In conventional experiments, human intervention is imposed on different settings, and then how subjects react to the intervention is recorded. In Glymour's approach, intervention is imposed on numbers, and how numbers react to the intervention is evaluated. But whether it is a true "intervention" is still debatable.

Woodward (2000, 2001) argued that a process or event could qualify as an intervention even if it does not involve human action. In other words, a purely "natural" process involving no animate beings at all can qualify as an intervention if causal information is embedded. This kind of research is often described by scientists as a "natural experiment." Moreover, even when manipulations are carried out by human beings, it is the causal features of those manipulations that matter for recognizing and characterizing causal relationships. For example, an intervention on variable X with respect to a second variable Y is a causal process that changes X in an appropriately exogenous way, so that if a change in Y occurs, it occurs only by virtue of the change in X, and not as a result of some other set of causal factors.

In experiments human intervention actually happens in the real world. In the mathematical realm, intervention or manipulation happens in a counterfactual fashion, or in the possible worlds. The intervention yields answers to questions like "what would happen to Y if X_1 were added to the model and the coefficient of X_2 were down-weighted?" In this case, whether or not the interventions that set the value of Xs and Y are carried out by human beings and whether or not they have in fact taken place is irrelevant (Hausman & Woodward, 1999).

Following this idea, intervention in the sense of TETRAD is legitimate. Human intervention in experiments does not create causal information or make the data ready for causal inferences. Causal properties have already been embedded in the subject matter and experimental control is just a way to reveal the causal information. If the data are non-experimental, causal characteristics are still within the data. Mathematical intervention, by the same token, is to make

the causal relationship more obvious, if there is any. In TETRAD, causal structure is represented in a system of equations. When the researcher changes the variables and/or the coefficients of the equations, he/she is changing the mechanism(s) or relationship(s) represented by it. Woodward (1999) stated that we can view this as a matter of intervening on the dependent variable in the equation so that the value of that variable is now fixed by the intervention rather than by the variables that previously determined its value.

Invariance

The second point made by Woodward on invariance is also relevant to Glymour's idea. It is important to note that in Woodward's framework, the concepts of invariance and interventions are closely related. For example, if Y=a+bX+e is a correct description of the causal relationship between X and Y, by intervening the value of X, the value of Y should change correspondingly within a reasonable range of data values. To be specific, in Woodward's view (2001) a causal generalization need not be universal. Rather, it could be just invariant, which means it is stable or robust in the sense that it would continue to hold under a relevant class of changes. This claim concerning causality is less ambitious than Cartwright's capacities, which aims to achieve context-free causation. Woodward (1998) asserted that many claims in social sciences are made within a limited range of circumstances. For example, interventions that change the money supply may change the price level in some range of circumstances, but not in others. Woodward (1999) and Hausman and Woodward (1999) gave a humorous example: If I water my plant with 1-3 liters of water, my plant would grow. If I water the same plant with 1,000 liters of water and the plant dies, it does not negate the statement "water causes plants to grow."

This simple argument provides a sound rationale for using linear modeling in TETRAD (Glymour, 1987). One popular argument against linear models is that it is too simplified to capture the complexity of the real world. In many situations, relationships between variables are non-linear. The relationship between stress and performance is a classical example. Psychologists found that for most people, as the stress level increases, the performance level increases correspondingly (see Figure 4). Nonetheless, the regression slope is reversed when too much pressure is imposed. Following Woodward's argument, if a professor assigned five term papers, six examinations, and ten presentations to her students, and as a result the whole class failed, is it right for her to say, "It seems that pressure does not improve performance in all situations. Now this causal law breaks down!" The main point here is that the validity of a causal model requires invariance within a

reasonable range of data values. Cases such as giving 1,000 liters of water to plants and six exams to a class must be dismissed.

Figure 4. Relationship between pressure and performance

Conclusion

In summary, in order to make causal interpretation of non-experimental data, the researcher must have some type of manipulation, rather than conditioning, of variables. The Causal Markov Condition and its sister, the common cause principle, provide the assumptions to structure relationships among variables in the path model and to load different variables into common latent constructs in the factor model. In addition, the Faithfulness Condition rules out those models in which statistical independence relations follow as a result of special coincidences among the parameter values. Putting all these together, TETRAD uses algorithms to examine all possible paths among variables to search for a plausible causal explanation. During this process intervention is imposed on the data and hence causal claims are justified.

Cartwright argued for "no causes in, no causes out," which means without background knowledge based upon empirical data, relevant variables and genuine causes may be omitted from the model. In this case, path searching and model building by sophisticated algorithms may be useless. The counter-arguments are that empirical data should not be narrowly defined in a restrictive sense, and the aim of the scientific investigation is not to examine all relevant variables. Cartwright objected to probabilistic causations and CMC by citing the Simpson's Paradox. However, statistical methodology is going toward broader generalizations in spite of the threat of the Simpson's paradox. Furthermore, Cartwright gave a counter-example to the faithfulness condition and warned that FC should not be adopted without careful investigation of the

background information. Nevertheless, rejecting FC may take the risk that any failure can be justified by adopting the notion that two equally powerful forces cancel out each other.

Woodward stated that even with non-experimental data, interventions are still possible because interventions are not necessarily carried out by humans at the stage of data collection. In data analysis, data values and equation parameters can be manipulated in a counterfactual manner. Moreover, the idea of invariance proposed by Woodward could be used to justify the linearity assumption of TETRAD.

Both Cartwright and Woodward made many other points rejecting or supporting TETRAD that could not be covered by this short paper. Additionally, besides Cartwright and Woodward, many other scholars from various disciplines, such as sociology, computer science, and mathematics, have participated in this type of discussion. This phenomenon shows that causality has become an inter-disciplinary subject matter. Even among philosophers discussion of this issue goes beyond pure philosophy. For example, although the primary role of Cartwright and Woodward is that of philosopher, Cartwright is versed in economics and Woodward has a mathematical background. Fruitful results are expected when input from such a wide variety of perspectives is integrated.

Acknowledgement

Special thanks to Professor Brad Armendt and Alex Novick for reviewing this paper.

References

Appleton, D. R. & French, J. M. (1996). Ignoring a covariate: An example of Simpson's paradox. American Statistician, 50, 340-341.

Cartwright, N. (1983). How the laws of physics lie. Oxford: Oxford University Press.

Cartwright, N. (1989). Nature' s capacities and their measurement. Oxford: Oxford University Press.

Cartwright, N. (1999a). Causal diversity and the Markov Condition. Synthese, 12, 3-27.

Cartwright, N. (1999b). The dappled world. Cambridge: Cambridge University Press.

Cartwright, N. (2001).What is wrong with Bayes Nets? Monist, 84, 242-276.

Data Description, Inc. (1999). DataDesk. [On-line] Available: http://www.datadesk.com

Druzdzel, M, & Glymour, C. (1995, March-April). Having the right tool: Causal graphs in

teaching research design. Paper presented at the University of Pittsburgh Teaching Excellence Conference, Pittsburgh, PA.

Dupres, J., & Cartwright, N. (1988). Probability and causality: Why Hume and indeterminism don't mix. Nous, 22, 521-536.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G. V., McGraw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage Publications.

Glymour, C. (1982). Casual inference and causal explanation. In R. McLaughlin (Ed.), What? Where? When? Why? Essays on induction, space, and time, explanation (pp. 179-191). Boston, MA: D. Reidel Publishing Company.

Glymour, C. (1983). Comments: Statistics and metaphysics. Journal of the American Statistical Association, 81, 964-966.

Glymour, C., Scheines, R., Spirtes, P, & Kelly, K. (1987). Discovering causal structure: Artificial intelligence, philosophy of science, and statistical modeling. Orlando, FL: Academic Press, Inc.

Glymour, G. (1999). Rabbit hunting. Synthese, 121, 55-78.

Glymour, C. & Cooper, G. F. (Eds.). (1999). Computation, causation, and discovery. Cambridge, MA: MIT Press.

Hausman, D., & Woodward, J. (1999). Independence, invariance, and the Causal Markov Condition. British Journal of Philosophy of Science, 50, 521-583.

Hunter, J. E., & Schmidt, F. L. (1990). Methods of meta-analysis: Correcting error and bias in research findings. Newbury Park, CA: Sage Publications.

Hsu, L. M. (1989). Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. Journal of Consulting and Clinical Psychology, 57, 131-137.

Meek, C., & Glymour, C. (1994). Conditioning and intervening. British Journal of Philosophy, 45, 1001-1021.

Olkin, I. (2000, November). Reconcilable differences: Gleaning insight from independent scientific studies. ASU Phi Beta Kappa Lecturer Program, Tempe, Arizona.

Pearl, J. (2000). Causality: Models, reasoning, and inference. New York: Cambridge University

Press.

Reichenbach, H. (1956). The direction of time. Berkeley, CA: University of Los Angeles Press.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society, Ser. B., 13, 238-241.

Spirtes, P., Glymour, C. & Scheines, R. (1993). Causation, Prediction and Search. Springer Verlag.

Ting, K. (1998). The TETRAD approach to model respecification. Multivariate Behavioral Research, 33, 157-164

Woodward, J. (1998). Causal independence and faithfulness. Multivariate Behavioral Research, 33, 129-148.

Woodward, J. (1999). Causal interpretation in systems of equations. Synthese, 121, 199-247.

Woodward, J. (2000). Explanation and invariance in the special sciences. British Journal of Philosophy of Science, 51, 197-254.

Woodward, J. (2001). Causation and manipulability. Stanford Encyclopedia of Philosophy. [On-line] Available: http://plato.stanford.edu/entries/causation-mani/

Yu, C. H. (2001). An introduction to computing and interpreting Cronbach Coefficient Alpha in SAS. Proceedings of 26th SAS User Group International Conference.

Yu, C. H., Behrens, J., & Ohlund, B. (under review). Abduction, deduction, and induction: Their applications in quantitative methods.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. Journal of Royal Statistical Society, 89, 1-69.