BOOK REVIEWS

Causality: Models, Reasoning, and Inference. By Judea Pearl. New York: Cambridge University Press, 2000. 384 pp. $40.00.

In *Causality: Models, Reasoning, and Inference*, Judea Pearl offers the methodological community a major statement on causal inquiry. His account of the philosophy and history of causal analysis is a joy to read, especially his spirited 28-page epilogue, "The Art and Science of Cause and Effect." The heart of Pearl's *Causality* is the set of ideas that he and his colleagues in computer science have developed over the past 20 years. They comprise a unified methodology for the graphical representation of joint probability distributions along with rules for inferring causality directly from such graphical representations.

Reading *Causality*, however, feels a bit (I would imagine) like going for a hike in a rain forest with Judea Pearl as your guide. Initially, one's excitement is aroused by first contact with his causal flora and then further enchanted by his claims of all that lies ahead. But, before long, one is so deeply enveloped in his graphical vegetation that fear of ever finding a way out begins to take over. As I write this review, I must admit that I am still circling about in the counterfactual mangroves of Chapters 7 through 10.

Nonetheless, I have come to a recommendation: Anyone who intends to teach causality in graduate methodology classes should read this book. Anyone who does not should avoid it, at least until the rest of us figure out which of its many ideas are unique and lasting contributions to causal analysis. That is to say, from the perspective of a sociologist, the work as a whole is a memorable tour de force, which succeeds admirably in provoking deep thoughts about (1) what social scientists can and should do with available observational data and (2) what sorts of theories and data sources we should be attempting to cultivate. But it is unclear which of Pearl's specific ideas should or

SOCIOLOGICAL METHODS & RESEARCH, Vol. 32, No. 3, February 2004 411-422

will uniquely alter the way in which empirical research is conducted in the social sciences.

The book unfolds as follows. After reviewing basic probability theory (and adopting a subjectivist variant of it), Pearl introduces directed graphs as a compact way of representing conditional independence restrictions for complex multidimensional probability distributions. He then comes to terms with his own past research, in which he now recognizes that he too closely aligned causality with probabilistic dependence. After settling these preliminaries, he defines the conditions necessary for the existence of a causal model, reinterprets the past 40 years of purported causal modeling, and develops specialized tools to deal with feedback effects, instrumental variables, and bounds for broken experiments.

In the course of presentation, Pearl covers, with his novel toolbox of concepts, much terrain that will be familiar to sociologists—structural equation models, Simpson's paradox, and counterfactual conditionals, for example. At times, his fresh perspective on these topics is easily grasped, as when familiar identification results are extended to more general nonparametric scenarios. And ironically, in these instances, it is not clear that it is worth the effort to learn all of the contours of the framework. At other times, however, Pearl reveals complexities that offer insight and that one cannot easily envision developing with the more familiar language of probability distributions or covariance algebra.

In general, the material is presented in formal but comprehensible doses. And some of the apparatus is quite charming. For example, the notion of a "collider" variable is introduced to explain how marginally independent probability distributions become dependent when one conditions on their common outcome. The common outcome variable is the collider, in the sense that it is the graphical terminus of prior causal variables. Other pieces of the framework, however, are less beguiling. For example, although it is indeed useful to think through formal identification puzzles without introducing error terms into the mix, as one can then avoid the rigidity of parametric models, I challenge any sociologist or economist to think through the fundamental Markov assumption for the existence of a causal model without introducing the idea of independence across error terms into one's own mind. Indeed, it is not surprising that Pearl himself gradually migrates toward the language of error terms when he adopts functional casual models to account for the complications introduced by confounders and concomitants.

The most memorable concept that Pearl develops is the do-operator, which offers a physical metaphor (in situations where *intervention* is too vague) for the ideas that are discussed customarily with the awkward language of counterfactuals. As all *Sociological Methods & Research* readers know, the conditional distribution(s) of Y given X (i.e., the conditional distribution that one can recover from the observable joint distribution of Y and X) does not necessarily coincide with the conditional distribution(s) of Y that one would obtain after setting X to alternative values through targeted interventions. In Pearl's syntax, this claim is conveyed simply as $\Pr(Y|X)$ does not necessarily equal $\Pr(Y|\,\mathrm{do}(X))$. Graph theory then supplies the necessary intuition (without having to introduce the potential outcomes and counterfactual conditionals that some scholars feel are unnecessarily metaphysical). When contemplating the distribution of Y given $\mathrm{do}(X)$, one simply ignores the variables that have determined X in the prevailing possible world that has generated the data. These extant "parents" of X are entirely ignorable, and thus when shifting to do-conditioning, the directed graph is "mutilated" by removing the directed edges from the parents of X to X itself.

At the risk of oversimplifying, it is not unfair to say that Pearl's definition of causality rests primarily on this distinction between $\Pr(Y|X)$ and $\Pr(Y|\,\mathrm{do}(X))$, thereby bringing the crux of his argument in line with the counterfactual model of causality. Pushing a good deal of formality aside (which is indeed crucial for understanding all that Pearl offers but is simply impossible to convey in a short review), a causal model for Pearl is a structural model in the structural equations tradition that perfectly matches the set of induced distributions that would emerge if one applied a do-operation successively to each variable in the graph. Accordingly, Pearl writes,

> The distinctive feature of causal models is that each variable is determined by a set of other variables through a relationship (called "mechanism") that remains *invariant* when those other variables are subjected to external influences. Only by virtue of this invariance do causal models allow us to predict the effect of changes and interventions, capitalizing on the locality of such changes. (P. 60)

The generic complication, of course, is that the measured variables in one's observed multidimensional probability distribution may not be the fundamental causal variables of interest. Thus, the set of relationships that one can observe may have an unknowable relationship to the true causal model that would be revealed by doing the variables of genuine interest. Even if one is willing (perhaps out of desperation) to specify a putative set of causal mechanisms by asserting conditional independencies for the data at hand, the useful factorization formulas that Pearl develops may not yield meaningful estimates of warranted causal effects. Pearl's inability, like everyone else, to offer us a way out of this common predicament is a genuine limitation of his framework, albeit one that perhaps reveals our own limitations as a discipline more than those of his book (as I discuss more below).

Nonetheless, one can evade Pearl's account of causality by adopting entirely different definitions of causes and effects and then arguing on principle that these represent the proper elements for foundational axioms of causal inquiry. Since Pearl has been justifying his framework for more than a decade, he seems to have developed a defense against all critics who take such a strategy. For example, in response to those who privilege explanatory rather than manipulative accounts of causality, Pearl writes,

> Explanatory accounts of causality . . . [maintain that] . . . causal models need not encode behavior under intervention but instead aim primarily to provide an "explanation" or "understanding" of how data are generated. Regardless of what use is eventually made of our "understanding" of things, we surely would prefer an understanding in terms of durable relationships, transportable across situations, over those based on transitory relationships. The sense of "comprehensibility" that accompanies an adequate explanation is a natural byproduct of the transportability of (and hence of our familiarity with) the causal relationships used in the explanation. . . . It thus seems reasonable to suggest that, in the final analysis, the explanatory account of causation is merely a variant of the manipulative account, albeit one where interventions are dormant. Accordingly, we may as well view our unsatiated quest for understanding "how data is generated" or "how things work" as a quest for acquiring the ability to make predictions under a wider range of circumstances, including circumstances in which things are taken apart, reconfigured, or undergo spontaneous change. (Pp. 25-26)

Through such argumentation, Pearl develops a relentless case for his graph-theoretic foundation for causal analysis. And since in most respects, it is consistent with the burgeoning literature on counterfactual causality, he offers an effective defense of that tradition as well.

This does not mean that Pearl's intellectual project will be unboundedly influential, as I suggested earlier. His focus on elemental events and the variables that encode them is indispensable for explication of his approach. But his resulting recipe for causal analysis, which works well in the closed systems that computer scientists can construct, is considerably harder to follow using the noisy and incomplete data with which social scientists work. This disjuncture will prevent *Causality* from having a major direct influence on empirical research practice in the social sciences.

But is this disjuncture really a limitation of Pearl's project, or rather is it a limitation of our own? Surely, Pearl is correct in implying that social scientists will never achieve the respect we crave until we (1) stop engaging in purported but invalid causal analysis but (2) somehow manage to discover a few genuine causal mechanisms. And thus, it may be that Pearl's work will be indirectly influential by further motivating us to craft a discipline for which his framework is more obviously relevant. In this regard, I am reminded of a memorable quotation of Otis Dudley Duncan, which he issued at the end of his introductory text on structural equation models (a book that, on its own, demands rereading after digesting Pearl, as Duncan had more of it right than the legacy of his work reveals):

> A strong possibility in any area of research at a given time is that there are *no* structural relations among the variables currently recognized and measured in that area. Hence, whatever its mathematical properties, no model describing covariation of those variables will be a structural model. What is needed under the circumstances is a theory that invents the proper variables. (Duncan 1975:152)

Perhaps the greatest lesson of the 28 years since Duncan offered this observation is how many areas of research in sociology were, and regrettably have remained, in such unproductive stasis.

If there is one immediate practical recommendation of Pearl's work, it is that we attempt to escape from our burden of bad data (and, at the same time, perhaps abandon most of the vacuous theorizing that is informed by it). It is hard to deny, after reading *Causality*, that social experiments and instrumental variables represent the most useful techniques for investigating causal mechanisms when independence of error terms across our favorite observed variables cannot be assumed (or, as Pearl would have it, when our variables cannot be encoded in a Markovian graph). In this way, Pearl's work offers yet more justification for the appeals of others that (1) we should enact social experiments whenever possible, and (2) when social experiments are impossible, we should design surveys to answer specific causal questions, trying at the same time to cultivate instrumental variables that can illuminate them.

In sum, *Causality* is a wonderful book. Pearl succeeds in bringing together in a general nonparametric framework the counterfactual tradition of causal analysis with the variants of structural equation modeling worth keeping. The graph theory that he uses to accomplish this fusion is often elegant. Thus, *Causality* is a major statement, which all who claim to know what causality is must read.

REFERENCES

Duncan, Otis Dudley. 1975. *Introduction to Structural Equation Models*. New York: Academic Press.

—*STEPHEN L. MORGAN,*
Cornell University

The Theory of the Design of Experiments. By D. R. Cox and N. Reid. Boca Raton, FL: Chapman & Hall/CRC Press, 2000. 323 pp. $74.95.

Courses and texts on experimental design often focus on the statistical analysis of data collected under given experimental arrangements, with primary attention devoted to the general linear model and elaborate analyses of variance and secondary attention (at best) given to the construction of designs to achieve desirable analytical goals. Cox and Reid reverse these orders of emphasis, stressing the role of design for the removal of systematic error and the reduction of random error.