



Statistics and Causal Inference: Comment: Statistics and Metaphysics

Clark Glymour

Journal of the American Statistical Association, Vol. 81, No. 396. (Dec., 1986), pp. 964-966.

Stable URL:

<http://links.jstor.org/sici?sici=0162-1459%28198612%2981%3A396%3C964%3ASACICS%3E2.0.CO%3B2-H>

Journal of the American Statistical Association is currently published by American Statistical Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/astata.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

- (1984), "Interaction" (with discussion), *International Statistical Review*, 52, 1–31.
- Cox, D. R., and Snell, E. J. (1981), *Applied Statistics*, London: Chapman & Hall.
- Pratt, J. W., and Schlaifer, R. (1984), "On the Nature and Discovery of

- Structure" (with discussion), *Journal of the American Statistical Association*, 79, 9–33.
- Wilk, M. B., and Kempthorne, O. (1957), "Non-additivities in the Latin Square Design," *Journal of the American Statistical Association*, 52, 218–236.

Comment

Statistics and Metaphysics

CLARK GLYMOUR*

1. INTRODUCTION

Holland's paper is as much philosophical analysis as it is statistics. The general lines of the account of causal relations he gives are familiar to philosophers, although he does not discuss any of the philosophical literature in which they may be found. I will try to place Holland's account in the framework of contemporary philosophical discussions of causality. I agree with the general thrust of his analysis, but I think certain restrictions he imposes are unwarranted, and I will say which they are, and why I think them unjustified.

Holland's account of causality is counterfactual. A fair paraphrase of his analysis is this:

Treatment t causes individual u to have the value Y_t for variable Y rather than the value Y_c for that variable if and only if u received treatment t , u has the value Y_t , and if u had received the treatment c rather than the treatment t , then u would have the value Y_c for variable Y .

Holland imposes conditions on this analysis, conditions that can be thought of as further explications of what he means it to say:

1. It must have been possible for u to have received treatment c rather than treatment t .
2. A treatment t can only be a cause of individual u having the value Y_t rather than Y_c provided t is a treatment that is applied to that same individual, u , and c is a treatment that could have been applied to that same individual.
3. Causation is a relation between two treatments and two possible variable states. The notion of t causing Y_t , without specification of any alternative treatment, or any alternative state of Y , is not defined.

I will consider these conditions later. First, I want to address the philosophical context.

2. COUNTERFACTUALS AND CAUSALITY

Notice that the clause following the phrase "if and only if" in my paraphrase of Holland's account is a counterfac-

tual conditional. It is a sentence of the form (neglecting tense):

If X were the case then Y would be the case.

Such sentences exhibit logical features that have interested philosophical logicians for some years. Their logical features include the following:

1. Counterfactuals can be logically false:

If X were the case then X and not X would be the case.

2. Counterfactuals can logically entail one another:

If X were the case then Y would be the case

entails

If X were the case then Y or Z would be the case.

3. Counterfactuals have different logical entailment relations than do ordinary material conditionals.

If X then Y

entails

If X and Z then Y ,

but

If X were the case then Y would be the case

does not entail

If X were the case and Z were the case then Y would be the case.

("If I had struck the match just now it would have lighted" is true, but "If I had struck the match just now and there had been no oxygen in the room, it would have lighted" is false.)

There are two principal ways to give a theory of the logical structure of some piece of reasoning. Both share the presupposition that the reasoning can be represented in a formalized language. One way is to characterize the logic axiomatically, by specifying an initial set of logical

*Clark Glymour is Professor and Head, Department of Philosophy, Carnegie-Mellon University, Pittsburgh, PA 15213.

truths and a set of rules of derivation, such that all and only the logical truths are derivable from the axioms, and such that if an inference is valid, then the conclusion of the inference is derivable from the premises of the inference. Another way to characterize logical structure is through formal semantics. A semantic characterization specifies a class of possible interpretations of the language and what it is for a sentence in the language to be true under an interpretation. The logical truths are then those that are true under every possible interpretation; valid arguments are those for which no interpretations exist making their premises true and their conclusion false. The aim of philosophical logicians studying some logical feature of language is to represent that feature in a formalized language, to characterize it both axiomatically and semantically, and to prove that the two characterizations determine exactly the same class of logical truths and the same collection of valid arguments.

There are two well-known logical theories of counterfactual conditionals, one due to Robert Stalnaker at Cornell University (Stalnaker 1984), the other to David Lewis (Lewis 1973, 1983), who is Holland's neighbor at Princeton University. There is also a logical theory of tensed counterfactual conditionals due to Richmond Thomason (Thomason and Gupta 1980). The Stalnaker and Lewis theories differ slightly, but the semantic characterization Stalnaker gives is especially simple, and I will, therefore, use it.

Imagine that there are a collection of possible worlds, much as in science fiction stories, and that in each possible world every sentence that is not counterfactual and is in our formalized language is either true or false. Further imagine that there is a relation between possible worlds, a relation of closeness. Finally, assume that for every possible world w and every logically possible condition A , there is a unique world that is the closest world to w in which A is true. If A happens to be true in w , then w itself is that closest possible world. Then for any sentences X and Y in the language,

"If X were the case then Y would be the case" is true in world w if and only if in the closest world to w in which X is true, Y is also true.

You may well wonder what possible worlds are and which possible worlds are supposed to be closer to which others. The point is that, for the purpose of giving a logical theory, it does not matter what possible worlds are or which of them are closest to which others. The worlds and their relations can be taken seriously or as a convenient mathematical fiction; in either case, they characterize the set of logical truths and they characterize valid inferences. At the very least, talk of possible worlds and their proximities provides a vivid metaphor that is easy to think about mathematically. At most, it provides the metaphysical underpinnings of our understanding of possibility and necessity.

Lewis proposed that causal relations are counterfactual relations. He proposed that if X and Y are sentences describing the occurrence of particular events, then

X causes Y if and only if X occurs and Y occurs and if X had not occurred then Y would not have occurred.

In the semantics of counterfactuals, this becomes

X causes Y if and only if X is true in the actual world and Y is true in the actual world and in the closest (to the actual world) possible world in which X is not true, Y is not true.

Return now to Holland's characterization of causal relations. We can see that his account is straightforwardly interpreted within the semantics of counterfactuals and that his account is really a specialization of Lewis's. Counterfactual analyses of causation, such as those of Lewis and of Holland, are naturally compared with alternative accounts that characterize causal relations in terms of probability relations. Such accounts have been provided, in various ways, by Suppes (1970), Granger (1969), Reichenbach (1949), Salmon (1980), and Skyrms (1980). Probabilistic accounts of causality have the advantage that they seem to make it easy to understand how we can have knowledge of causal relations, and equally, to ease our understanding of the bearing of statistics on causal inference. Technical details aside, causal inference becomes a statistical estimation problem. They have the disadvantage that they do not always accord very well with our intuitive judgments about causal relations.

Counterfactual accounts of causality have the disadvantage that they appeal to unobservables—to what would be true if . . . , and to what goes on in possible worlds we will never see. They, therefore, present us with a mystery as to how we can know anything about causal relations. The mystery surely has a solution, and the general lines of the solution must be something like this: We are able to infer causal relations because we are able to infer counterfactual truths, and we are able to infer counterfactual truths because we make assumptions that we test against one another in rather indirect ways. Holland's article seems to me especially valuable in clarifying some of these assumptions and in explicating their relations. The philosophical community, unfortunately, has not been very energetic in addressing the mystery.

3. HOLLAND'S RESTRICTIONS

I am not convinced that the restrictions Holland imposes on causal relations are equally justified. Consider, first, the requirement that for treatment t to cause individual u to have Y_t rather than Y_c it must have been possible for u to have received treatment c instead of t . Holland intends this requirement to exclude factors such as genetic constitution and attributes determined by genetic constitution (e.g., race and gender) from the category of causes. There is no treatment that would give one and the same individual a genetic structure other than the actual one. There seem to be two ideas here. One is that genetic structure is not an event, not a happening but an enduring attribute, and causes

must be events. The other is that the identity of organisms depends on their genetic structure, so any actual or possible individual who differs from me in genetic structure is not me. Thus counterfactuals whose antecedents suppose that I had a different genetic structure are nonsensical.

If we insist that only events, not attributes, can be causes, then we can still make sense of the talk of causal attributes as a *façon de parler*. We need only find for each individual and each attribute the event that was the acquisition of that attribute by that individual. In the case of genetic structure there is such an event, conception. In many of the sociological cases in which attributes are used as causes and that Holland rejects as meaningless, there are also appropriate events that are the acquisition of the attributes, and the talk of attributes as causes can, therefore, be interpreted as a harmless convenience of speech. I cannot agree that "The causal model literature has not been careful in separating meaningful and meaningless causal statements and path diagrams" (p. 958). There is little need for this sort of care.

We can identify persons across at least some alterations in genetic structure. Down's syndrome is caused by a trisomy—a bit of extra genetic material attached to a chromosome pair. If that extra bit of material were removed from the zygote, without damaging viability, the zygote would develop into a person—the very same person I should say—without Down's syndrome. Even when one cannot identify persons across changes in genetic structure, there may still be correspondences that make counterfactuals intelligible. My parents tell me that if I had been a girl, I would have been named Olga. I believe what they tell me, and I think they mean more by it than that their intent was to name their first-born "Olga" if their first-born was female. (I believe this because I believe they did things like the following: referring to the creature in my mother's uterus, they said, "If it's a girl, we will call her Olga." The reference was not just to whatever person should be their first-born, but, as it turns out, to me, and the antecedent of the conditional is contrary to fact.) I can imagine a possible world in which I do not exist, but a female counterpart of me does. In that world she is conceived on the day I was conceived in this world, her parents in that world are mine in this, and her name is Olga. If counterparts are conceivable—and why not?—then counterfactuals that violate identity conditions are intelligible, and if counterfactuals are intelligible, then causal relations are as well.

Holland's second restriction is that the treatment that is to be called a cause must be applied to the very individual that has the variable value that is called the effect. I see no clear motivation for this restriction, and it certainly does not agree with our causal judgments and knowledge. The Big Bang caused the cosmological background radiation. A parent's acquisition of syphilis can cause a child's (congenital) syphilis, and so forth. Nothing in the counterfac-

tual analysis of causation requires such a restriction, and I am rather at a loss to find a motive for its introduction.

I am tempted to think that Holland's third restriction, which demands that a cause always be relative to a specific alternative, is an improvement on the bare counterfactual account of causal relations. The reason is this: My Uncle Schlomo smoked two packs of cigarettes a day, and I am firmly convinced that smoking two packs of cigarettes a day caused him to get lung cancer. But it may not be true that in the closest possible world in which Uncle Schlomo did not smoke two packs a day, he did not contract cancer. Reflecting on Schlomo's addictive personality, and his general weakness of will, it may well be that the closest possible world in which Schlomo did not smoke two packs of cigarettes a day is a world in which he smoked three packs a day. I can reconcile this reflection with the counterfactual analysis of causality by supposing, with Holland, that "smoking two packs of cigarettes a day caused him to get lung cancer" is elliptical speech, and what is meant, but not said, is that smoking two packs of cigarettes a day, rather than not smoking at all, caused Schlomo to contract lung cancer.

4. CONCLUSION

Probability may have begun with games of chance, but one of the principal goals of statistics has always been the determination of causal relations from both experimental and nonexperimental data. I applaud Holland's willingness to try to make the links a little clearer, and I even agree in the main with what I take to be his understanding of causal relations. I applaud as well his efforts to connect philosophy and statistics. Statistics runs with a lot of philosophy, too much of it tacit, and bad philosophy is best avoided by explicitness. I would only caution against branding discourse that does not agree with a philosophical account as "meaningless." People talk as they will, and if they talk in a way that does not fit some piece of philosophical analysis and seem to understand one another well enough when they do, then there is something going on that the analysis has not caught. That is not a failing of the speakers. It is, if anything, a failing of we who philosophize, even if we philosophize with statistics.

ADDITIONAL REFERENCES

- Lewis, D. (1973), *Counterfactuals*, Cambridge, MA: Harvard University Press.
- (1983), *Philosophical Papers* (Vol. 2), Oxford, U.K.: Oxford University Press.
- Reichenbach, H. (1949), *The Theory of Probability*, Berkeley and Los Angeles: University of California Press.
- Salmon, W. (1980), "Probabilistic Causality," *Pacific Philosophical Quarterly*, 61, 50–74.
- Skyrms, B. (1980), *Causal Necessity*, New Haven, CT: Yale University Press.
- Stalnaker, R. (1984), *Inquiry*, Boston, MA: Bradford Books.
- Thomason, R., and Gupta, A. (1980), "A Theory of Conditionals in the Context of Branching Time," *Philosophical Review*, 88, 65–90.