Comment

D. R. COX*

It is a pleasure to have the chance of congratulating Holland on an exceptionally lucid article on an important topic. Indeed the issues explicitly and implicitly raised by the article seem to me more important for the foundations of our subject than discussion of the nature of probability, perennially intriguing though that may be. Philosophy is often regarded by scientists, on this side of the Atlantic at least, as an irredeemably "soft" subject, but here is a matter both of philosophical interest and also with important practical implications, for example, for the interpretation of coefficients in multiple regression equations.

The question of what can constitute a cause in this context is a key issue, and there is need for some good terminology. Cox and Snell (1981, p. 12) called variables that, in the context under consideration, should not be regarded as treatments, intrinsic variables. A subdivision into those associated with the individual person, animal, or whatever and those associated with the environment is sometimes useful. It might also be useful to distinguish between treatments and quasitreatments. In addition, the term nonspecific (Cox 1984) may be used for strata, blocks, and so forth that are normally intrinsic, but with no clearly specified unique characterization.

The notion that certain variables cannot properly be regarded as causes is most concisely encapsulated in the physicists' notion that passage of time cannot be regarded as a cause: of course, a process going on in time, such as molecular rearrangement, could be a cause, because it is possible to conceive of time passing without the rearrangement in question taking place.

In some observational studies the distinction between quasitreatment and intrinsic variables is a matter of viewpoint and may not be clear-cut. Think, for example, of an observational study of alcohol consumption related to some outcome variable.

One point deserving emphasis is the need for careful specification of what constitutes a particular treatment, including what may be subsidiary consequences. This may be crucial if technically correct but nevertheless misleading conclusions are to be avoided. In studying the effect of alcohol, is diet held fixed?

A celebrated, if possibly apocryphal, example concerns an agricultural field trial in which one treatment led to such a superior quality and quantity of product that birds for many kilometers around gathered to consume the product, leading to poor final yield. Does that treatment cause poor yield? In one sense, yes. Similar points arise in clinical trials in connection with the distinction between intention to treat and per protocol analyses. The point partly is that the difficulties of observational studies cannot be totally avoided in randomized experiments, if one is to look in depth at interpretation. The searching discussion of Pratt and Schlaifer (1984) is very relevant.

This is related to the issue of "layers" of interpretation. Is not the reason that one expects turning a light switch to have the result it does not just direct empirical observation but a subtle and deep web of observations and ideas—the practice of electrical engineering, the theory of electrical engineering, various ideas in classical physics, summarized, in particular, in Maxwell's equations, and underneath that even ideas of unified field theory? One reason that the notion of "cause" is so important is that it carries suggestions of relations at a deeper level of interpretation than the direct observation under study.

My final comment concerns absence of interaction or presence of unit-treatment additivity. Holland suggests in Section 6 of his article that this might have been regarded as a "technical" requirement, whatever that might mean. In fact it seems to me to be of great importance from various points of view. First the condition is not wholly operationally verifiable, as Holland carefully discusses. A rigid adherent of operationalism might, therefore, regard the condition as meaningless; in fact, so far as I can see, rigid operationalism went out of favor a long time ago, both in philosophy with the decline of logical positivism and in physics with increased emphasis on quantum mechanics. Yet it represents a fine ideal, that all assumptions and concepts should be capable of direct verification, but in the present context, and in many others, partial operationalism seems to be the most one can reasonably get. This is that certain aspects of the assumption can be tested.

Thus in the present context one could detect use of an inappropriate scale, or, as soon as intrinsic variables are available, examination for treatment $\times$ intrinsic interaction becomes feasible. Such considerations are important both for understanding and for examining possible extrapolation of the conclusions to new units. When no such further information is available the technical questions raised by Neyman for the Latin square remain (Wilk and Kempthorne 1957); that is, is the usual analysis unbiased? I think it is arguable that the analysis is unbiased in a reasonable sense (Cox 1958), but admittedly a somewhat contorted view of the question under study is needed.

In conclusion, I welcome the article as an account of underdiscussed issues of considerable importance.

ADDITIONAL REFERENCES

Cox, D. R. (1958), "The Interpretation of the Effects of Non-additivity in the Latin Square," *Biometrika*, 45, 69–73.

* D. R. Cox is Professor of Statistics, Department of Mathematics, Imperial College, London SW72BZ, England.

Journal of the American Statistical Association
December 1986, Vol. 81, No. 396, Theory and Methods

——— (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 counterfactual 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.

Journal of the American Statistical Association
December 1986, Vol. 81, No. 396, Theory and Methods