Forthcoming: Observational Studies
Special Issue on William Cochran's contributions.

TECHNICAL REPORT
R-456
July 2015

# Causal Thinking in the Twilight Zone

Judea Pearl
University of California, Los Angeles
Computer Science Department
Los Angeles, CA, 90095-1596, USA
(310) 825-3243 / judea@cs.ucla.edu

To students of causality, the writings of William Cochran provide an excellent and intriguing vantage point for studying how statistics, lacking the necessary mathematical tools, managed nevertheless to cope with increasing demands for policy evaluation from observational studies. Cochran met this challenge in the years 1955-1980, when statistics was preparing for a profound, albeit tortuous transition from a science of data, to a science of data generating processes. The former, governed by Fisher's dictum (1922) "the object of statistical methods is the reduction of data" was served well by the traditional language of probability theory. The latter, on the other hand, seeking causal effects and policy recommendations, required an extension of probability theory to facilitate mathematical representations of generating processes.

No such representation was allowed into respectable statistical circles in the 1950-60s, when Cochran started looking into the social effects of public housing in Baltimore. While data showed improvement in health and well-being of families that moved from slums to public housing, it soon became obvious that the estimated improvement was strongly biased; Cochran reasoned that in order to become eligible for public housing the parent of a family may have to possess both initiative and some determination in dealing with the bureaucracy, thus making their families more likely to obtain better healthcare than non-eligible families.[1] This led him to suggest "adjustment for covariates" for the explicit purpose of reducing this causal effect bias. While there were others before Cochran who applied adjustment for various purposes, Cochran is credited for introducing this technique to statistics (Salsburg, 2002) primarily because he popularized the method and taxonomized it by purpose of usage.

Unlike most of his contemporaries, who considered cause-effect relationships "ill-defined" outside the confines of Fisherian experiments, Cochran had no qualm admitting that he sought such relationships in observational studies. He in fact went as far as *defining* the objective of an observational study: "to elucidate causal-and-effect relationships" in situations where controlled experiments are infeasible (Cochran, 1965). Indeed, in the paper before us, the word "cause" is used fairly freely, and other causal terms such as "effect," "influence," and "explanation" are almost as frequent as "regression" or "variance." Still, Cochran was well aware that he was dealing with unchartered extra-statistical territory and cautioned us:

[1]Narrated in Cochran (1983, p. 24).

> "Claim of proof of cause and effect must carry with it an explanation of the mechanism by which this effect is produced."

Today, when an analyst declares that a claim depends on "the mechanism by which an effect is produced" we expect the analyst to specify what features of the mechanism would make the claim valid. For example, when Rosenbaum and Rubin (1983) claimed that propensity score methods may lead to unbiased estimates of causal effects, they conditioned the claim on a counterfactual assumption named "strong ignorability." Such identifying assumptions, though cognitively formidable, provided a formal instrument for proving that some adjustments can yield unbiased estimates. Similarly, when a structural analyst makes the claim that an "indirect effect" is estimable from observational studies, the claim must follow assumptions about the structure of the underlying graph which, again, assures us of zero-bias estimates (see Pearl (2014b)).

Things were quite different in Cochran's era; an appeal to "a mechanism," like an appeal to "subject matter information" stood literally for a confession of helplessness, since "mechanisms" and causal relationships had no representation in statistics. Structural equation models (SEM), the language used by economists to represent mechanisms, were deeply mistrusted by statisticians, who could not bring themselves to distinguish structural from regression models (Guttman, 1977; Freedman, 1987; Cliff, 1983; Wermuth, 1992; Holland, 1995).[2] Counterfactuals, on the other hand, were still in the embryonic state that Neyman left them in – symbols with no model, no formal connection to realizable variables, and no inferential machinery with which to support or refute claims.[3] Fisher's celebrated advice: "make your theories elaborate" was no help in this transitional era of pre-formal causation; there is no way to elaborate on a theory that cannot be represented in some language.

It is not surprising, therefore, that Cochran's conclusions are quite gloomy:

> "It is well known that evidence of a relationship between $x$ and $y$ is no proof that $x$ causes $y$. The scientific philosophers to whom we might turn for expert guidance on this tricky issue are a disappointment. Almost unanimously and with evident delight they throw the idea of cause and effect overboard. As the statistical study of relationships has become more sophisticated, the statistician might admit, however, that his point of view is not very different, even if he wishes to retain the terms cause and effect."

It is likewise not surprising that in the present article, Cochran does not offer readers any advice on which covariates are likely to reduce bias and which would amplify bias. Any such advice, as we know today, requires a picture of reality, which Cochran understood to be both needed and lacking at his time.[4] On the positive side, though, he did have the vision to anticipate the emergence of a new type of research paradigm within statistics, a paradigm centered on mechanisms:

[2]This mistrust persists to some degree even in our century, see Berk (2004) or Sobel (2008).

[3]These had to wait for Rubin (1974), Robins (1986), and the structural semantics of Balke and Pearl (1994).

[4]To the best of my knowledge, the only adjustment-related advice in the entire statistics literature prior to 1980 was Cox's warning that "the concomitant observations be quite unaffected by the treatments" (Cox, 1958, p. 48); an exceptional defiance of an unwritten taboo against the use of data-generating models.

> "A claim of proof of cause and effect must carry with it an explanation of the mechanism by which the effect is produced. Except in cases where the mechanism is obvious and undisputed, this may require a completely different type of research from the observational study that is being summarized."

I believe the type of research we see flourishing today, based on a symbiosis between the graphical and counterfactual languages (Morgan and Winship, 2014; VanderWeele, 2015; Bareinboim and Pearl, 2015) would perfectly meet Cochran's vision of a "completely different type of research." This research differs fundamentally from the type of research conducted in Cochran's generation. First, it commences with a commitment to understanding what reality must be like for a statistical routine to succeed and, second, it represents reality in terms of data-generating models (read: "mechanisms"), rather than probability distributions.

Encoded as nonparametric structural equations, these models have led to a fruitful symbiosis between graphs and counterfactuals and have unified the potential outcome framework of Neyman, Rubin, and Robins with the econometric tradition of Haavelmo, Marschak, and Heckman. In this symbiosis, counterfactuals (potential outcomes) emerge as natural byproducts of structural equations and serve to formally articulate research questions of interest. Graphical models, on the other hand, are used to encode scientific assumptions in a qualitative (i.e., nonparametric) and transparent language and to identify the logical ramifications of these assumptions, in particular their testable implications.[5]

A summary of results emerging from this symbiotic methodology is given in Pearl (2014a) and includes complete solutions[6] to several long-standing problem areas, ranging from policy evaluation (Tian and Shpitser, 2010) and selection bias (Bareinboim et al., 2014) to external validity (Bareinboim and Pearl, 2015; Pearl and Bareinboim, 2014) and missing data (Mohan et al., 2013).

This development has not met with universal acceptance. Cox and Wermuth (2015), for example, are still reluctant to endorse the tools that this symbiosis has spawned, questioning in essence whether interventions cannot be mathematized.[7] Others regard the symbiosis as unscientific (Rubin, 2008) or less than helpful (Imbens and Rubin, 2015, p. 22), insisting for example that investigators should handle ignorability judgments by unaided intuition.

I strongly believe, however, and I say it with a deep sense of responsibility, that future explorations of observational studies will rise above these inertial barriers and take full advantage of the tools that the graphical-counterfactual symbiosis now offers.

---

[5] Note that the potential outcome framework alone does not meet these qualifications. Scientific assumptions must be converted to conditional ignorability statements (Rosenbaum and Rubin, 1983; Imbens and Rubin, 2015) which, being cognitively formidable, escape the scrutiny of plausibility judgment and impede the search for their testable implications.

[6] By "complete solution" I mean a method of producing consistent estimates of (causal) parameters of interests, applicable to any hypothesized model, and accompanied by a proof that no other method can do better except by strengthening the model assumptions.

[7] Unwittingly, the very calculus that they reject happens to resolve the problem that they pose ("indirect confounding") in just four steps (Pearl, 2015a, http://www.mii.ucla.edu/causality/; Pearl, 2015b).

# References


Balke, A. and Pearl, J. (1994). Counterfactual probabilities: Computational methods, bounds, and applications. In *Uncertainty in Artificial Intelligence 10* (R. L. de Mantaras and D. Poole, eds.). Morgan Kaufmann, San Mateo, CA, 46–54.

Bareinboim, E. and Pearl, J. (2015). Causal inference from big data: Theoretical foundations and the data-fusion problem. Tech. Rep. R-450, <http://ftp.cs.ucla.edu/pub/stat_ser/r450.pdf>, Department of Computer Science, University of California, Los Angeles, CA. Forthcoming, *Proceedings of the National Academy of Sciences*.

Bareinboim, E., Tian, J. and Pearl, J. (2014). Recovering from selection bias in causal and statistical inference. In *Proceedings of the Twenty-eighth AAAI Conference on Artificial Intelligence* (C. E. Brodley and P. Stone, eds.). AAAI Press, Palo Alto, CA. Best Paper Award, <http://ftp.cs.ucla.edu/pub/stat_ser/r425.pdf>.

Berk, R. (2004). *Regression Analysis: A Constructive Critique*. Sage, Thousand Oaks, CA.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research* **18** 115–126.

Cochran, W. (1965). The planning of observational studies of human population. *Journal of the Royal Statistical Society (Series A)* **128** 234–255.

Cochran, W. G. (1983). *Planning and Analysis of Observational Studies*. Wiley, New York.

Cox, D. (1958). *The Planning of Experiments*. John Wiley and Sons, NY.

Cox, D. and Wermuth, N. (2015). Design and interpretation of studies: Relevant concepts from the past and some extensions. *Observational Studies,* to appear.
URL `http://arxiv.org/pdf/1505.02452v1.pdf`

Fisher, R. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London, Series A* **222** 311.

Freedman, D. (1987). As others see us: A case study in path analysis (with discussion). *Journal of Educational Statistics* **12** 101–223.

Guttman (1977). What is not what in statistics. *The Statistician* **26** 81–107.

Holland, P. (1995). Some reflections on Freedman's critiques. *Foundations of Science* **1** 50–57.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, New York, NY.

MOHAN, K., PEARL, J. and TIAN, J. (2013). Graphical models for inference with missing data. In *Advances in Neural Information Processing Systems 26* (C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K. Weinberger, eds.). Curran Associates, Inc., 1277–1285. URL `http://papers.nips.cc/paper/4899-graphical-models-for-inference-with-missing-data.`

MORGAN, S. L. and WINSHIP, C. (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research.* 2nd ed. Cambridge University Press, New York, NY.

PEARL, J. (2014a). The deductive approach to causal inference. *Journal of Causal Inference* **2** 115–129.

PEARL, J. (2014b). Interpretation and identification of causal mediation. *Psychological Methods* **19** 459–481.

PEARL, J. (2015a). Indirect Confounding and Causal Calculus (On three papers by Cox and Wermuth). Blog entry: http://www.mii.ucla.edu/causality/.

PEARL, J. (2015b). Indirect Confounding and Causal Calculus (On three papers by Cox and Wermuth). Tech. Rep. R-457, <http://ftp.cs.ucla.edu/pub/stat_ser/r457.pdf>, Department of Computer Science, University of California, Los Angeles, CA.

PEARL, J. and BAREINBOIM, E. (2014). External validity: From *do*-calculus to transportability across populations. *Statistical Science* **29** 579–595.

ROBINS, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period – applications to control of the healthy workers survivor effect. *Mathematical Modeling* **7** 1393–1512.

ROSENBAUM, P. and RUBIN, D. (1983). The central role of propensity score in observational studies for causal effects. *Biometrika* **70** 41–55.

RUBIN, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66** 688–701.

RUBIN, D. (2008). Author's reply (to Ian Shrier's Letter to the Editor). *Statistics in Medicine* **27** 2741–2742.

SALSBURG, D. (2002). *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century.* Henry Holt and Company, LLC, New York, NY.

SOBEL, M. (2008). Identification of causal parameters in randomized studies with mediating variables. *Journal of Educational and Behavioral Statistics* **33** 230–231.

TIAN, J. and SHPITSER, I. (2010). On identifying causal effects. In *Heuristics, Probability and Causality: A Tribute to Judea Pearl* (R. Dechter, H. Geffner and J. Halpern, eds.). College Publications, UK, 415–444.

VANDERWEELE, T. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction.* Oxford University Press, New York, NY.

Wermuth, N. (1992). On block-recursive regression equations. *Brazilian Journal of Probability and Statistics* (with discussion) **6** 1–56.