

Comments on Woodward’s article

Judea Pearl

Cognitive Systems Laboratory
Computer Science Department
University of California, Los Angeles, CA 90024
judea@cs.ucla.edu

As much as I appreciate Woodward’s drive to attribute deeper meaning to the notion of “intervention,” I cannot convince myself that this simple concept should be mystified by complex definitions such as SIN and W/SIN. As a devout interventionist, I have found unchallenged comfort and clarity in the standard “equation wipe-out” conception, where “intervening on X ” means removing the equation that determines X and replacing it with some other equation (often a constant). In *Causality* (2000), I have shown that this ideal conception, dubbed “atomic,” can serve as a basis for formulating and analyzing more complex types of interventions, for example, actions that change several mechanisms at once, or actions that merely modify, not wipe-out the mechanism controlling X .

In Section 2, Woodward argues that the notion of “hypothetical idealized intervention” is needed as a theoretical tool for understanding the meaning of structural equations. This is indeed a worthwhile endeavor, one that enjoys a simple characterization in the atomic framework. For example, the meaning of Eq. 3 in Woodward’s example is straightforward: If we were to intervene and set X_2 to x_2 and X_3 to x_3 , then the value y of Y would be $y = ax_2 + bx_3$, regardless of the value we set for X_1 .¹

In Section 3, on the other hand, Woodward seems to abandon the *idealized* character of our basic target, and seeks conditions under which *practical*, yet non-ideal interventions would meet certain (informally stated) desiderata. He then posits these intuitive conditions as a definition for our target notion of “hypothetical idealized intervention.”

This strategy stands contrary to common scientific practice. Normally, we define the meaning of scientific terms (e.g., basic physical quantities) using ideal experiments and, when faced with non-ideal experiments, we ask whether they would still allow us to deduce our target quantities, namely, quantities that would be directly measured under ideal experiments. Woodward himself recognizes the wisdom of this practice, stating (Section 6): “On my view, the content

¹Formally, this reads: $Y(\text{set}(x_1, x_2, x_3)) = ax_2 + bx_3$, where $Y(\text{set}(x))$ stands for the counterfactual “ Y , had X been x ”.

of causal claims often may be clarified by invoking what would happen under hypothetical interventions even if those interventions cannot or will not be carried out.” Thus, I fail to see why Woodward does not take the “atomic intervention” as a primitive and, when faced with some non-ideal manipulation, derive (not merely posit) conditions such as SIN, which are sufficient for answering questions of interest, say whether X has a causal influence on Y .

In *Causality* (2000) I demonstrate the ease with which this approach can be executed. For example, Section 3.4.4 (page 88) gives conditions under which the causal effect of X on Y can be deduced from an experiment in which another variable, Z , is randomized instead of X . These conditions include SIN as a very special case. To tackle the more general experiment where a manipulated variable, I , modifies several mechanisms, we simply add node I to the graph and draw arrows from I to all variables determined by the modified mechanisms. Subsequently, if our aim is find the causal effect of X on Y , we ask whether the target quantity $P(y|\text{set}(x))$ can be deduced from $P(y, x, z, w, \dots | \text{set}(I))$ – the joint distribution obtained under manipulations of I .² More modestly, if our aim is merely to verify whether X has causal influence on Y , an aim that seems to be at the center of Woodward’s concerns, we ask whether the truth value of the proposition “ $P(y|\text{set}(x)) = P(y)$ ” can be deduced from $P(y, x, z, w, \dots | \text{set}(I))$.

I should also note that Woodward’s assertion (Section 10) that the “atomic” conception of intervention requires that “we already have some information about the causal relationship, if any, between X and Y ” is incorrect; Knowing whether Z affects the causal relationship between X and Y does not require any prior information about that relationship. It requires merely an assumption on whether Z may change Y if we hold X fixed (i.e., whether an arrow should be drawn between Z and Y), an assumption invoked in SIN as well.

In summary, I have found that, invariably, questions about interventions and experimentation, ideal as well as non-ideal, practical questions as well as philosophical, can be formulated precisely and managed systematically using the atomic intervention as a primitive notion. Based on Woodward’s discussion, I infer that Nancy Cartwright too has been having difficulties with notions of intervention, modularity and invariance. I will thus end this commentary with a daring conjecture (or a challenge) that any difficulty which Woodward or Cartwright can articulate unambiguously can be expressed formally in the language of atomic interventions and reduced to a mathematical problem in the calculus of $P(y|\text{set}(x))$.

References

[Pearl, 2000] J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, 2000.

²Formal machinery for deducing such quantities is provided in chapter 3 of *Causality* (2000).