THEORIA

On the unity between observational and experimental causal discovery

(Sobre la unidad entre el descubrimiento causal observacional y experimental)

Jiji ZHANG*
Hong Kong Baptist University

ABSTRACT: In "Flagpoles anyone? Causal and explanatory asymmetries", James Woodward supplements his celebrated interventionist account of causation and explanation with a set of new ideas about causal and explanatory asymmetries, which he extracts from some cutting-edge methods for causal discovery from observational data. Among other things, Woodward draws interesting connections between observational causal discovery and interventionist themes that are inspired in the first place by experimental causal discovery, alluding to a sort of unity between observational and experimental causal discovery. In this paper, I make explicit what I take to be the implicated unity. Like experimental causal discovery, observational causal discovery also relies on interventions (or exogenous variations, to be more accurate), albeit interventions that are not carried out by investigators and hence need to be detected as part of the inference. The observational patterns appealed to in observational causal discovery are not only surrogates for *would-be* interventions, as Woodward sometimes puts it; they also serve to mark relevant interventions that *actually* happen in the data generating process.

KEYWORDS: causal discovery; exogenous variation; intervention; interventionism; invariance; observational data.

RESUMEN: En "Flagpoles anyone? Causal and explanatory asymmetries", James Woodward complementa su celebrada teoría intervencionista de la causación y la explicación con nuevas ideas sobre asimetrías causales y explicativas, extraídas de recientes métodos de descubrimiento causal a partir de datos observacionales. Entre otras cosas, Woodward establece interesantes conexiones entre el descubrimiento causal observacional e ideas intervencionistas inspiradas inicialmente en el descubrimiento causal experimental, aludiendo a cierta unidad entre el descubrimiento causal observacional y experimental. Al igual que el descubrimiento causal experimental, el descubrimiento causal observacional también se apoya en intervenciones (o variaciones exógenas, para ser más precisos), aunque sean intervenciones que no son realizadas por investigadores y por tanto tienen que ser detectadas como parte de la inferencia. Los patrones observacionales a los que se apela en el descubrimiento causal observacional no son los sustitutos de posibles intervenciones, como Woodward algunas veces sugiere; también sirven para marcar intervenciones relevantes que de hecho tienen lugar en el proceso de generación de datos.

PALABRAS CLAVE: descubrimiento causal; variaciones exógenas; intervención; intervencionismo; invariancia; datos observacionales.

* **Correspondence to:** Jiji Zhang. Department of Religion and Philosophy, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong – zhangjiji@hkbu.edu.hk – https://orcid.org/0000-0003-0684-2084

1. Introduction

For several decades now, how to discover causal relations between variables using statistical methods has been a vigorous research program pursued in several fields. One of the main goals is to invent principled and reliable ways to infer which variable has a (direct) causal influence on or is (directly) causally relevant to which variable in a multivariate system, from observational data and without prior knowledge or assumption about the causal order. The term 'observational' is used to indicate the absence of any active control or manipulation by investigators of the data generating process under investigation; observational data are, so to speak, generated by the system of interest running its natural course. By contrast, experimental data are generated by a process that includes active control or manipulation by investigators. A paradigmatic example is a randomized controlled trial, where the allocation of treatments is designed and administered by the investigators (with a randomization scheme). Causal inference based on experimental data is regarded as by and large more reliable than that based on observational data, but the importance and potential of the latter are getting increasingly acknowledged and appreciated, due on the one hand to the relative abundance of observational data, especially in the era of big data, and on the other hand to significant methodological advances in recent years (Peters *et al.,* 2017).

James Woodward's rich and illuminating article (Woodward, 2022) amply demonstrates that the cutting-edge methods for causal discovery from observational data (or observational causal discovery as I will henceforth call it) have novel implications for philosophical theorizing about causal and explanatory asymmetries. A champion of the influential interventionist approach to causation and explanation, Woodward draws interesting connections between observational causal discovery and interventionist themes that are inspired in the first place by experimental causal discovery, alluding to, as I understand it, a sort of unity between observational and experimental causal discovery. In this paper I follow up on this issue and make the implicated unity more explicit. My main thesis is that like experimental causal discovery, observational causal discovery, when applicable, also relies on interventions (or exogenous variations, to be more accurate), albeit interventions that are not carried out by investigators and hence need to be identified as part of the inference. Observational causal discovery is epistemologically more challenging in large part because of the additional need to infer the loci of interventions.

To proceed, I will first review in Section 2 some distinctive features of Woodward's (2003) notion of an intervention, and two ways in which this notion is used to characterize the presence of a causal relation. I suggest that the essential element in Woodward's notion of intervention is a notion of exogenous variation, and that although Woodward prefers not to build a condition of invariance or mechanism-preservation into the notion of an intervention, the kind of intervention or exogenous variation that matters in causal inference must satisfy some condition of invariance. This makes salient the possibility of sometimes detecting exogenous variation through invariance, which is related to Woodward's latest discussion of a value-relationship independence/invariance principle for causal discovery. Then, in Section 3, I briefly recall Richard Scheines's (2005) argument that the inference to the presence of a causal relation using observational conditional dependence and independence relations is essentially the same as the inference based on experimental data. I suggest that the essential common element is that they both infer the presence of a causal relation between two variables from their covariation in which one variable's variation is

(known or assumed or inferred to be) exogenous with respect to the other. I apply this idea, in Section 4, to examine the more recent and powerful methods of observational causal discovery discussed by Woodward (2022), suggesting subtle modifications to some of his interpretations while endorsing his main points. I close in Section 5 with brief concluding remarks.

2. Woodward's conception of intervention

In his seminal work, Woodward (2003, p. 98) presented a careful definition of what counts as an intervention on one variable with respect to another. For present purposes, we need not go into more details than noting a few features of the account. First, Woodward explicitly relativizes an intervention on a variable with respect to another variable, so we talk about an intervention on variable X with respect to variable Y rather than an intervention on X simpliciter. Second, an intervention on X with respect to Y is represented as an intervention variable taking a value, and the core requirement for an intervention variable for X with respect to Y is that it influences Y, if at all, only through X and is statistically independent of all other variables that influence Y without going through X.[1] Third, and very important for my thesis, an intervention does not necessarily involve a human action; any garden-variety variable may serve as an intervention variable for X with respect to Y as long as it stands in the right causal relations with X and Y. Fourth, an intervention is not required to be mechanism-preserving; that is, it is possible that an intervention on X with respect to Y changes how Y is affected by X (though by definition, this effect of the intervention cannot be a result of the intervention affecting causes of Y whose influences do not go through X.)

The restriction to singletons in the first feature is not essential. Woodward's definition can be easily extended to cover interventions on a set of variables with respect to another set of variables, but for his main purposes the version for singletons is sufficient. The third feature Woodward refers to as *nonanthropomorphism*. A straightforward implication of this feature is that even though no intervention is carried out by investigators in an observational study, an intervention on a variable of interest with respect to another may nonetheless have taken place. Therefore, it is at least coherent to say that observational causal discovery also relies on interventions.

The fourth feature is probably the most distinctive and controversial. As is noted by Woodward, many other influential accounts of interventions or manipulations, such as those of Spirtes *et al.* (2000) and Pearl (2009), build in some version of a mechanism-preservation or invariance condition. Woodward's main worry regarding those accounts is that

---

[1] In Woodward (2003), another requirement put down for an intervention variable for X is that there are some possible values of the variable such that when the variable takes these values, X ceases to depend on its other causes and depends solely on the intervention variable. These values implement what are sometimes referred to as "hard" or "surgical" interventions. However, in the subsequent definition of an intervention, there does not seem to be any requirement that these values be taken by the intervention variable in an intervention. In any case, Woodward (2022) also talks about "soft" intervention variables, so I will assume in this paper that this requirement is not imposed on an intervention variable.

the notion of intervention on X with respect to Y would then already invoke the causal relationship between X and Y, and to use such a notion as he does to characterize the causal relationship between X and Y would smell of a potentially objectionable kind of circularity. This is an interesting point, but more relevant to my present purpose is an apparent difficulty with the more liberal notion of an intervention. To see the difficulty, let us compare two ways a causal relation between X and Y may be characterized in Woodward's framework.

One way is to say (roughly) that

(1) X is a (type-level) cause of Y if and only if there is an intervention on X with respect to Y that would change the value of X, under which the value (or probability distribution) of Y would also change.

Woodward (2003, p. 108) considered the possibility that if an intervention is not required to be mechanism-preserving, then some interventions may end up destroying the causal influence of X on Y and not changing the value of Y. As he rightly pointed out, this possibility does not threaten (1), whose right hand side is existentially quantified. However, I worry about the possibility of a "false positive": an intervention on X with respect to Y that is not mechanism-preserving may change the value of Y, in which case X is declared a cause of Y by (1), but intuitively a mechanism-altering intervention is not a good test of the *pre-intervention* causal relation between X and Y. This issue becomes more salient if we compare (1) to another way of characterizing the causal relationship between X and Y, which goes through a notion of invariance. Although Woodward does not require every intervention on X with respect to Y to preserve whatever causal mechanism there is between X and Y, invariance under *some* such interventions is regarded as a necessary condition for a generalization relating X and Y to be causal (with the causal direction going from X to Y).[2] Since, as I think it is safe to assume, X is a (type-level) cause of Y only if they are related by a true generalization that is causal (with the direction from X to Y), we can also say that

(2) X is a (type-level) cause of Y only if there is a true generalization relating X and Y and an intervention on X with respect to Y such that the generalization remains true under the intervention.

Call an instance that makes an existentially quantified statement true a *witnessing* instance. A witnessing intervention for the right hand side of (2) is required to satisfy a kind of invariance or mechanism-preservation, whereas a witnessing intervention for that of (1) is apparently not. Given how central the notion of invariance is in Woodward's account of causal and explanatory generalizations, this apparent discrepancy between (1) and (2) should probably be resolved or accounted for in favor of (2). This does not mean that (1) must be extensionally inadequate without an explicit requirement of invariance, because it may be argued that only a mechanism-preserving intervention can satisfy the prop-

---

2 There is a small technical catch in regarding it as a sufficient condition, and Woodward introduces a further qualification called a *testing intervention* to render the necessary condition also sufficient (see also Woodward and Hitchcock, 2003). Since I can make my point without entering that complication, I will just work with an "only if" statement here.

erty stated in the right hand side of (1), or that whenever there is a witnessing intervention that is not mechanism-preserving, there is also a witnessing intervention that is. But it suggests strongly that the kind of intervention that matters for causal inference in the spirit of (1) needs to satisfy some condition of mechanism-preservation. Moreover, although the additional requirement often amounts to just another assumption, it also creates the opportunity of sometimes detecting the presence of an intervention by checking invariance or some surrogate property of a hypothesized causal generalization. We will return to this point in Section 4.

Finally, if we look back at the second feature noted earlier of Woodward's characterization of an intervention, together with the characterization of a causal relation between X and Y given in (1), it is clear that the relevant role of an intervention on X with respect to Y in the context of causal inference is to create a change or variation in X that is not due to a change in Y nor associated with any variation in a cause of Y whose influence on Y does not go through X. Call such a variation an *exogenous* variation of X with respect to Y. I suspect that most if not all of the insights of interventionism can be recast in terms of a notion of exogenous variation. I will not explore here how feasible and desirable it is to reformulate interventionist theories in terms of exogenous variation, but a potential advantage is worth noting: the notion of exogenous variation, unlike the notion of intervention-induced variation, accommodates the possibility of spontaneous and uncaused variations of a variable that are nonetheless exogenous with respect to another variable and are therefore as good as intervention-induced variations for the purpose of probing the causal relation between the variables.

This advantage is perhaps merely theoretical. It is plausible to think that in practical causal inference, exogenous variations always result from some interventions. Even so, as I will stress later, observational causal discovery relies heavily on detecting interventions (as opposed to interventions known in advance, as in experimental causal discovery), but in many cases, no intervention variable is explicitly located, and it is, strictly speaking, only the exogeneity of variation that is detected.

3. Scheines on the similarity between observational and experimental causal discovery

In addition to Woodward (2022), another inspiration for the main thesis of this paper is the insight of Scheines (2005). Scheines compared the basic rationale of experimental inference of a causal relation between two random variables and that of observational inference of a causal relation based on graphical modelling, where a directed graph is taken to represent both a causal structure and a statistical model defined by a set of statistical conditional independence constraints (Pearl, 1988, 2009; Spirtes *et al.,* 2000). The similarity between them highlighted by Scheines is depicted in Figure 1. In Figure 1(a), the presence of a causal influence of X on Y (represented by the arrow from X to Y) is inferred based on an experimental intervention on X, where the probability distribution of X is known (or assumed) to be determined by an intervention variable (say, a randomizing device) I_X, and the statistical dependence between X and Y under this intervention is taken to be evidence for the causal influence of X on Y. In Figure 1(b), the presence of a causal influence of X on Y is inferred from passive observations on X and Y as well as some other variables such as Z_1 and Z_2. Suppose the observations turn out to warrant the statement that Z_1 is statistically

independent of Z_2 (conditional on the empty set) and the statement that they are statistically independent of Y conditional on X, together with a number of conditional dependence statements.[3] Under two assumptions known as the causal Markov condition and the Faithfulness or Stability condition (Spirtes *et al.,* 2000; Pearl, 2009), it follows from these conditional independence and dependence statements that X does not have a causal influence on either Z_1 or Z_2 (as represented by the arrowheads at X on the edges between Z_1 and X, and between Z_2 and X), and that X has a causal influence on Y (as represented by the arrow from X to Y).

Scheines calls a variable such as Z_i (either Z_1 or Z_2) in Figure 1(b) a *detectible instrument* (for X with respect to Y) and stresses the following similarity between such a detectible instrument and the intervention variable I_X in Figure 1(a): both are adjacent to X but not adjacent to Y in the respective causal graph, and neither is influenced by X. As Scheines takes some care to explain, having a variable with these features allows an inference from certain observed statistical relations to the presence of a causal arrow from X to Y. That is why he calls such a variable an instrument (for causal inference concerning X and Y). In observational causal discovery such a variable is not known in advance to be an instrument but its status as an instrument is sometimes detectible from data, as the simple example in Figure 1(b) illustrates.
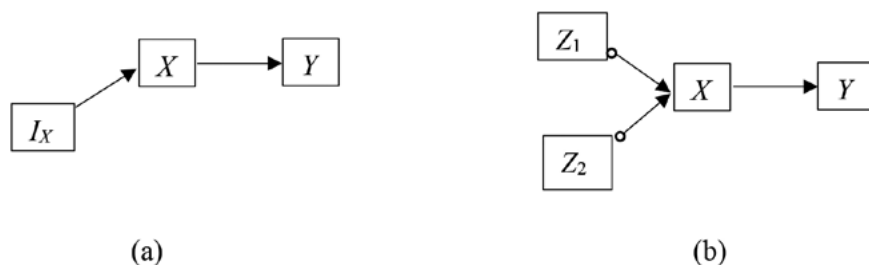

Figure 1

Scheines's (2005) comparison between experimental causal discovery and (a prominent approach to) observational causal discovery: (a) experimental inference to a causal arrow from X to Y based on investigators' interventions on X with respect to Y; (b) observational inference to a causal arrow from X to Y based on detecting Z_1 or Z_2 as an instrument.

A difference between Z_i in Figure 1(b) and I_X in Figure 1(a) is also noted by Scheines. The former is not necessarily a cause of X, as indicated by the circle at Z_i on the edge between Z_i and X,[4] whereas the latter is a cause of X. In other words, Z_i could be but is not

---

[3]   Namely, Z_1 and X are dependent conditional on any subset of the other variables, and so are Z_2 and X, and are X and Y, which is why each of these pairs of variables has an edge between them in Figure 1(b).

[4]   In the framework Scheines (2005) works with, $Z_i \circ^* X$ means roughly that it could be $Z_i \,{}^* X$, in which case Z_i is a cause of X, or $Z_i \ll X$, in which case Z_i is not a cause of X but there is a common cause of Z_i and X.

necessarily a (soft) intervention variable for X with respect to Y.[5] One way to view the role of Z_i in this simple example is that even if it is not itself an intervention variable for X with respect to Y, it is a surrogate for an unobserved intervention variable for X with respect to Y (namely, an unobserved common cause of X and Z_i). From this perspective, the example illustrates the idea of detecting interventions in observational causal inference, interventions that are not carried out by investigators. Another, closely related but on my view more apt conception is that the detected features of Z_i serve to show that some variation of X is exogenous with respect to Y, and the accompanying covariation of Y implies that X has a causal influence on Y. As explained by Scheines, the inference in this case is not just that X is a cause of Y, but moreover that there is no unobserved common cause of X and Y (see also the notion of "visibility" in Zhang, 2008). Since this simple example does not contain any observed common cause either, it is, so to speak, the whole variation of X that is detected to be exogenous with respect to Y. In more complex cases, it will be the variation of X conditional on or adjusted for some observed variables that is detectibly exogenous with respect to Y. Such exogenous variations are perhaps most plausibly interpreted as resulting from certain (soft) interventions on X with respect to Y, but conceptually they do not have to be. In any case, only a surrogate for an intervention variable may be detected in such inferences.

4. More recent advances in observational causal discovery

The kind of observational causal discovery discussed by Scheines (2005) remains a major approach to inferring causal structures from observational data. However, since 2006, a plethora of other methods for observational causal discovery have been proposed and refined, some of which are more powerful at least in the following respect: under suitable assumptions, they can reliably infer the causal relation between two variables without appealing to any detectible instrument in Scheines's sense. These more recent advances in observational causal discovery figure centrally in Woodward's (2022) new accounts of causal and explanatory asymmetries. I submit that although these methods do not proceed by identifying an observed variable as an intervention variable or a surrogate thereof, the detection of exogenous variations still plays a pivotal role in them.

Consider first the class of methods that employ the setup of noisy functional causal models. For simplicity, I will follow Woodward (and Scheines) to focus attention on the task of inferring the causal relation between two variables X and Y. The basic setup for many of these methods is to assume that the effect variable is a function of the cause variable and a noise or error term that is statistically independent of the cause variable. Thus, assuming without loss of generality that the causal arrow goes from X to Y, then it is assumed that the causal generalization relating X and Y can be properly represented as $Y = f(X, N)$, for some function f (of a certain form) and some noise term N (of a certain type) such that

---

[5] To repeat the clarification stated in footnote 1, I am following Woodward (2022) to allow intervention variables to be "soft", in that they are not required to relieve the target variable of dependence on other causes, but only required to satisfy the other conditions in Woodward's (2003) definition of an intervention variable.

X and N are statistically independent. As is observed by Woodward (2022, p. 34), one consequence of such an assumption is that the noise term N can be viewed as a (soft) intervention variable for Y with respect to X. This way of thinking of the noise term is warranted if the noise term is interpreted as representing omitted causes of Y, an interpretation that seems commonly adopted in practice. But again, even if the noise term is not interpreted as a cause of Y, but is instead regarded as representing a genuinely stochastic component in the generation of a value of Y from a value of X (Steel, 2005), it can still be viewed as picking out a part of the variation of Y that is exogenous with respect to X.

More important for my present purpose is the following point alluded to by Woodward. If the causal mechanism between X and Y is correctly represented as $Y = f(X, N)$, with X and N being statistically independent, then whatever generate the values of X "operate so as to change the value of X in a way that is independent of the other causes of Y" (Woodward, 2022, p. 34), and so may be regarded as interventions on X with respect to Y, or at any rate as exogenous variations of X with respect to Y (even if they are somehow spontaneous and uncaused). From this viewpoint, it is useful to think of the fitting of a noisy functional causal model as potential evidence for detecting interventions or exogenous variations: if data support the hypothesis of a functional relationship $Y = f(X, N)$ with statistical independence between X and N, then it *may* be evidence that the observed variation of X is exogenous with respect to Y.

It is not necessarily evidence because it is sometimes possible to fit such a model come what may, even when X does not cause Y. The best known example is that if X and Y follow a bivariate Gaussian distribution, then it always fits $Y = f(X, N)$ for some linear function f and some Gaussian noise N that is independent of X. On the other hand, if no restriction is put down on f or N, then it is always possible to fit a noisy functional causal model for any two random variables with continuous support (Hyvärinen and Pajunen, 1999; Zhang *et al.,* 2014). However, under some restrictions, such as linear non-Gaussian acyclic models (LiNGAM, Shimizu *et al.,* 2006), additive noise models (ANM, Hoyer *et al.,* 2009), or post-nonlinear models (PNL, Zhang and Hyvärinen, 2009), it becomes nontrivial to satisfy the requirement of statistical independence between the noise term and the hypothesized cause. It has been shown that either always (for LiNGAM) or generically (for ANM and PNL), this requirement can be met for at most one causal direction: if the joint probability distribution of X and Y is compatible with a generalization $Y = f(X, N)$ satisfying the model restrictions such that X and N are statistically independent, then it is not compatible with any generalization $X = g(Y, N')$ satisfying the model restrictions such that Y and N' are statistically independent. Moreover, it is clear that for LiNGAM models, if X and Y are confounded by an unobserved common cause, then under a suitable faithfulness assumption, the requirement of statistical independence between the noise term and the hypothesized cause can be met in neither direction (Entner and Hoyer, 2010). I suspect that this is also generically the case for ANM and PNL models.

With such model restrictions, a standard procedure to infer a causal relation from an observed covariation between X and Y is based on checking whether a noisy functional causal model (satisfying the model restrictions) is warranted by data in either direction. If it is warranted in one direction but not in the other, then the former direction is inferred to be the causal direction. Such a procedure is usually taken to assume in the first place that there is no unobserved confounding and the task is to decide between two hypotheses: that X causes Y (without confounding) or that Y causes X (without confounding). However,

we may also think of the procedure as trying to detect interventions or exogenous variations, using the asymmetric possibility of fitting a suitable model as a criterion. Viewed this way, the procedure may or may not return an informative answer, depending on whether the criterion is or is not met for judging a variable to have an exogenous variation with respect to the other variable in their covariation. If an uninformative answer or suspension of judgement is allowed, it is unnecessary to assume away latent confounding and force a selection between two hypotheses of an unconfounded causal relation. When we cannot fit a noisy functional causal model in either direction, an option is to remain silent about the causal relation between X and Y (and suggest that there is probably a latent confounder), because no suitable exogenous variation is detected.

Moreover, even when there is confounding and the variation of either variable is not entirely exogenous with respect to the other, it may be possible to identify parts of the variations that are exogenous and exploit them to draw informative causal inference. As mentioned earlier, the setup of noisy functional causal models renders the noise terms a sort of intervention variables or proxies for exogenous variations. Sometimes they may be recoverable from data to a sufficient extent even when there is latent confounding, and be used to infer causal relations. A case in point is the approach to inferring LiNGAM models based on independent component analysis (ICA). Even when latent confounding is present, how observed variables depend on the exogenous noise terms can be sufficiently recovered under a faithfulness assumption using the so-called overcomplete ICA, so that the acyclic causal structure among the observed variables can still be inferred (Hoyer *et al.,* 2008; Salehkaleybar *et al.,* 2020).

In addition to the methods based on noisy functional causal models, Woodward (2022) discussed a class of methods that exemplify what he calls the value-relationship independence/invariance (VRI) principle. As I understand it, the VRI principle is a further elaboration or enrichment of Woodward's characterization of causal and explanatory generalizations in terms of invariance under interventions (as we emphatically reviewed in Section 2). An important new element is that invariance under interventions can sometimes be indicated by or inferred from a sort of "independence" between variations in hypothesized cause variables and the hypothesized generalization relating the cause variables to an effect variable. The latter, in turn, may be inferred from observational data in various ways. Consider again the basic task of inferring the causal relation between two variables X and Y for illustration. The idea is that a hypothesized generalization relating X to Y, either represented by a functional relationship $Y = f(X)$ (as in, e.g., Janzing *et al.,* 2012) or represented by the conditional probability function $P(Y|X)$ (as in, e.g., Zhang *et al.,* 2015), may be judged to be "independent" of variations of X, according to various criteria of "independence" that can be checked based on observational data. This "independence", according to Woodward, is a "surrogate for or indicator of" the notion of invariance under interventions.

I find this viewpoint compelling, but would like to highlight a point that is not sufficiently stressed in Woodward's discussion. In the cases where we detect the desired "independence" between variations of X and a hypothesized generalization relating X to Y (and no such "independence" in the other direction), and on this basis infer that X is a cause of Y, we are also committed to inferring that whatever generates the observed values of X amounts to a Woodwardian intervention on X with respect to Y, or at any rate constitutes exogenous variations of X with respect to Y (even if they are somehow spontaneous

or uncaused). Therefore, the detected "independence" not only indicates *counterfactual* invariance under *world-be* interventions, as Woodward (2022) rightly claims, it also indicates that the *actual* data generating process features some interventions on X with respect to Y and that the generalization relating X to Y is *actually* invariant under these interventions.

This perspective of detecting interventions or exogenous variations through criteria of independence/invariance is more conspicuous in causal inference based on observational data from multiple populations or regimes, which feature different probability distributions but are assumed to share a causal structure (e.g., Peters *et al.,* 2016; Huang *et al.,* 2020; Mooij *et al.,* 2020). Woodward (2022) gives a simplified example of this sort on pp. 36-37, where $P(Y|X)$ is observed to be invariant in two datasets that feature different marginal distributions of X. As he describes in a footnote, one way to view this kind of inference is that the observed invariance across the two datasets is taken to indicate that the change in $P(X)$ results from an intervention on X with respect to Y, which, together with the accompanying change in $P(Y)$, licence the conclusion that X is a cause of Y. This criterion of invariance of one module in a causal network under changes of another module can be generalized into a criterion of independence between changes of different modules in a causal network (Huang *et al.,* 2020). The generalized criterion is akin to Woodward's (2022) brief discussion of the independence between different causal relationships (cf. Peters *et al.,* 2017), and can be used to detect where multiple interventions have taken place respectively in a complex data generating process.

5. Conclusion

According to Woodward's *non-anthropomorphic* interventionism, intervening is a matter of causing in the right way and may happen without an agent's deliberate design or action. Consequently, the data generating process in observational causal discovery, despite the absence of deliberate control by the investigator, often features relevant interventions that may be detected by various clever means. Following the lead of Scheines (2005) and Woodward (2020), I have tried to argue that a variety of approaches to observational causal discovery can be seen as attempts, in one way or another, to detect interventions or exogenous variations in the data generating process. On my view, therefore, observational causal discovery and experimental causal discovery are unified in at least the following sense: they both need detectible covariation under exogenous variation to draw a positive causal conclusion. In experimental studies, the exogenous variation is ensured by the experimental setup, if designed adequately, and only the covariation needs to be detected. In observational studies, however, the biggest challenge is to find ways to detect exogenous variations as well as the accompanying covariations.

This additional challenge for observational causal discovery is of course epistemologically significant. Relevant interventions are not guaranteed to be present in observational settings. Even when they are, it is not always possible under acceptable or reasonable assumptions to detect them. And even when it is possible, the assumptions that enable the detection, though plausible, are usually less secure than those needed to ensure the presence of an intervention in an experimental setting. It is therefore right to exercise more caution towards results of observational causal discovery, and when feasible, to validate the results

with experiments. Nonetheless, the unity between observational and experimental causal discovery, if real, suggests that the former's epistemological status is probably more continuous with the latter's than it is often thought to be.

REFERENCES

Entner, D. & Hoyer, P. O. (2010). Discovering unconfounded causal relationships using linear nongaussian models. *JSAI International Symposium on Artificial Intelligence* (pp. 181-195).

Hoyer, P. O., Shimizu, S., Kerminen, A. J., & Palviainen, M. (2008). Estimation of causal effects using linear non-gaussian causal models with hidden variables. *International Journal of Approximate Reasoning*, 49(2), 362-378.

Hoyer, P. O., Janzing, D., Mooij, J., Peters, J., & Schölkopf, B. (2009). Nonlinear causal discovery with additive noise models. *Advances in Neural Information Processing Systems* (pp. 689-696).

Huang, B., Zhang, K., Zhang, J., Ramsey, J., Sanchez-Romero, R., Glymour, C., & Schölkopf, B. (2020). Causal discovery from heterogeneous/nonstationary data. *Journal of Machine Learning Research*, 21(89), 1-53.

Hyvärinen, A. & Pajunen, P. (1999). Nonlinear independent component analysis: existence and uniqueness results. *Neural Networks*, 12, 429-439.

Janzing, D., Mooij, J., Zhang, K., Lemeire, J., Zscheischler, J., Daniušis, P., Steudel, B., & Schölkopf, B. (2012). Information-geometric approach to inferring causal directions. *Artificial Intelligence*, 182, 1-31.

Mooij, J., Magliacane, S., & Claassen, T. (2020). Joint causal inference from multiple contexts. *Journal of Machine Learning Research*, 21, 1-108.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Morgan Kaufmann.

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge: Cambridge University Press.

Peters, J., Bühlmann, P., & Meinshausen, N. (2016). Causal inference using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society, Series B*, 78(5), 947-1012.

Peters, J., Janzing, D., & Schölkopf, B. (2017). *Elements of causal inference: Foundations and learning algorithms.* Cambridge, MA: MIT Press.

Salehkaleybar, S., Ghassami, A., Kiyavash, N., & Zhang, K. (2020). Learning linear non-Gaussian causal models in the presence of latent variables. *Journal of Machine Learning Research*, 21, 1-24.

Scheines, R. (2005). The similarity of causal inference in experimental and non-experimental studies. *Philosophy of Science*, 72, 927-940.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., & Kerminen, A. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7, 2003-2030.

Spirtes, P., Glymour, G., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge, MA: MIT Press.

Steel, D. (2005). Indeterminism and the causal Markov condition. *The British Journal for the Philosophy of Science*, 56, 3-26.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford and New York: Oxford University Press.

Woodward, J. (2022). Flagpoles anyone? Causal and explanatory asymmetries. *THEORIA. An International Journal for Theory, History and Foundations of Science,* 37(1), 7-52 (https://doi.org/10.1387/theoria.21921).

Woodward, J. & Hitchcock, C. (2003). Explanatory generalizations, part I: A counterfactual account. *Noûs*, 37, 1-24.

Zhang, J. (2008). Causal reasoning with ancestral graphs. *Journal of Machine Learning Research*, 9, 1437-1474.

Zhang, K. & Hyvärinen, A. (2009). On the identifiability of the post-nonlinear causal model. *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence* (pp. 647-655).

Zhang, K., Wang, Z., Zhang, J., & Schölkopf, B. (2015). On estimation of functional causal models: general Results and application to the post-nonlinear causal model. *ACM Transactions on Intelligent Systems and Technology*, 7(2), 13:1-13:22.

Zhang, K., Zhang, J., & Schölkopf, B. (2015). Distinguishing cause from effect based on exogeneity. *Proceedings of the 15th Conference on Theoretical Aspects of Rationality and Knowledge* (pp. 261-271).

JIJI ZHANG is a professor in the Department of Religion and Philosophy at Hong Kong Baptist University. He is especially interested in the epistemological, logical, and methodological questions about causal inference.

ADDRESS: Department of Religion and Philosophy, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong
Email: zhangjiji@hkbu.edu.hk
ORCID: 0000-0003-0684-2084