Interventions and Causal Inference

Frederick Eberhardt and Richard Scheines†‡

The literature on causal discovery has focused on interventions that involve randomly assigning values to a single variable. But such a randomized intervention is not the only possibility, nor is it always optimal. In some cases it is impossible or it would be unethical to perform such an intervention. We provide an account of 'hard' and 'soft' interventions and discuss what they can contribute to causal discovery. We also describe how the choice of the optimal intervention(s) depends heavily on the particular experimental setup and the assumptions that can be made.

1. Introduction. Interventions have taken a prominent role in recent philosophical literature on causation, in particular in work by Woodward (2003), Hitchcock (2007), Cartwright (2002, 2006), and Hausman and Woodward (1999, 2004). Their work builds on a graphical representation of causal systems developed by computer scientists, philosophers and statisticians called "Causal Bayes Nets" (Pearl 2000; Spirtes, Glymour, and Scheines 2000). The framework makes interventions explicit, and introduces two assumptions to connect qualitative causal structure to sets of probability distributions: the Causal Markov and Faithfulness assumptions. In his recent book, *Making Things Happen* (2003), Woodward attempts to build a full theory of causation on top of a theory of interventions. In Woodward's theory, roughly, one variable X is a *direct cause* of another variable Y if there exists an intervention on X such that if all other variables are held fixed at some value, X and Y are associated. Such an account assumes a lot about the sort of intervention needed, however, and Woodward goes to great lengths to make the idea clear. For example, the intervention must make its target independent of its other causes, and it must directly influence only its target, both of which are ideas difficult to make clear without resorting to the notion of direct causation.

†To contact the authors, please write to: Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213; e-mail: fde@cmu.edu and scheines@cmu.edu.

‡The first author is funded by the Causal Learning Collaborative Initiative supported by the James S. McDonnell Foundation. Many aspects of this paper were inspired by discussions with members of the collaborative.

Philosophy of Science, 74 (December 2007) pp. 981–995. 0031-8248/2007/7405-0036$10.00
Copyright 2007 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

Statisticians have long relied on intervention to ground causal inference. In *The Design of Experiments* (1935), Fisher considers one treatment variable (the purported cause) and one or more effect variables (the purported effects). This approach has since been extended to include multiple treatment and effect variables in experimental designs such as Latin and Graeco-Latin squares and Factor experiments. In all such cases, however, one must designate certain variables as potential causes (the treatment variables) and others as potential effects (the outcome variables), and inference begins with a randomized assignment (intervention) of the potential cause.

Although randomized trials have become de facto the gold standard for causal discovery in the natural and behavioral sciences, without such an a priori designation of causes and effects Fisher's theory is far from complete. First, without knowing ahead of time which variables are potential causes and which effects, more than one experiment is required to identify the causal structure—but we have no account of an optimal sequence of experiments. Second, we don't know how to statistically combine the results of two experiments involving interventions on different sets of variables. Third, randomized assignments of treatment are but one kind of intervention. Others might be more powerful epistemologically, or cheaper to execute, or less invasive ethically, etc.

The work we present here describes two sorts of interventions ('structural' and 'parametric') that seem crucial to causal discovery. These two types of interventions form opposite ends of a whole continuum of 'harder' to 'softer' interventions. The distinction lines up with interventions of different "dependency," as presented by Korb et al. (2004). We then investigate the epistemological power of each type of intervention without assuming that we can designate ahead of time the set of potential causes and effects. We give results about what can and cannot be learned about the causal structure of the world from these kinds of interventions, and how many experiments it takes to do so.

2. Causal Discovery with Interventions. Causal discovery using interventions depends not only on what kind of interventions one can use, but also on what kind of assumptions one can make about the models considered. We assume that the causal Markov and faithfulness assumptions are satisfied (see Cartwright 2001, 2002 and Sober 2001 for exceptions and Steel 2005 and Hoover 2003 for sample responses). The causal Markov condition amounts to assuming that the probability distribution over the variables in a causal graph factors according to the graph as so:

$$\mathrm{P}(\boldsymbol{X}) = \prod_{X_i \in \boldsymbol{X}} \mathrm{P}(X_i | \mathrm{parents}(X_i)),$$

https://doi.org/10.1086/525638 Published online by Cambridge University Press

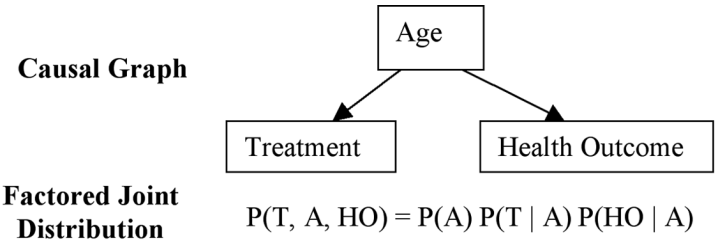


Figure 1.—A causal graph.

where the parents of X_i are the immediate causes of X_i, and if there are no parents then the marginal distribution $P(X_i)$ is used. For example, consider Figure 1.

The faithfulness assumption says that the causal graph represents the causal independence relations true in the population, that is, that there are no two causal paths that cancel each other out. However, several other conditions have to be specified in order to formulate a causal discovery problem precisely.

It is often assumed that the causal models under consideration satisfy causal sufficiency and acyclicity. Causal sufficiency is the assumption that there are no unmeasured common causes of any pair of variables that are under consideration (no latent confounders). Assuming causal sufficiency is unrealistic in most cases, since we rarely measure all common causes of all of our pairs of variables. However, the assumption has a large effect on a discovery procedure, since it reduces the space of models under consideration substantially. Assuming acyclicity prohibits models with feedback. While this may also be an unrealistic assumption—in many situations in nature there are feedback cycles—this is beyond the scope of this paper.

With these assumptions in place we turn to 'structural' and 'parametric' interventions. The typical intervention in medical trials is a 'structural' randomization of one variable. Subjects are *assigned* to the treatment or control group based on a random procedure, for example, a flip of a coin. We call the intervention "structural" because it alone completely determines the probability distribution of the target variable. It makes the intervened variable independent of its other causes and therefore changes the causal structure of a system before and after such an intervention. In a drug trial, for example, a fair coin determines whether each subject will be assigned one drug or another or a placebo. The assignment of treatment is independent of any other factor that might cause the outcome.[1] Ran-

1. For example, treatment and age should not be associated, as age is a likely cause of response to almost any drug.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

domization of this kind ensures that, at least in the probability distribution true of the population, such a situation does not arise.

In Causal Bayes Nets a *structural* intervention is represented as an exogenous variable I (a variable without causes) with two states (*on*/*off*) and a single arrow into the variable it manipulates.[2] When I is *off*, the passive observational distribution obtains over the variables. When I is *on*, all other arrows incident on the intervened variable are removed, and the probability distribution over the intervened variable is a determinate function of the intervention only. This property underlies the terminology "structural."[3] If there are multiple simultaneous structural interventions on variables in the graph, the manipulated distribution for each intervened variable is independent of every other manipulated distribution,[4] and the edge breaking process is applied separately to each variable. This implies that *all* edges between variables that are subject to an intervention are removed. After removing all edges from the original graph incident to variables that are the target of a structural intervention, the resulting graph is called the *postmanipulation graph* and represents what is called the manipulated distribution over the variables.

More formally, we have the following definition for a *structural* intervention I_s on a variable X in a system of variables $\boldsymbol{V}$ (see Figure 2):

- I_s is a variable with two states (*on*/*off*).
- When I_s is *off*, the passive observational distribution over $\boldsymbol{V}$ obtains.
- I_s is a direct cause of X and only X.
- I_s is exogenous,[5] that is, uncaused.
- When I_s is *on*, I_s makes X independent of its causes in $\boldsymbol{V}$ (breaks the edges that are incident on X) and *determines* the distribution of X; that is, in the factored joint distribution $\mathrm{P}(\boldsymbol{V})$, the term $\mathrm{P}(X|\mathrm{parents}(X))$ is replaced with the term $\mathrm{P}(X|I_s)$, all other terms in the factorized joint distribution are unchanged.

2. For more detail, see Spirtes, Glymour, and Scheines 2000, Chapter 7; see also the appendix of this paper.

3. Pearl (2000) refers to these interventions as 'surgical' for the same reason. They are also sometimes referred to as 'hard' or 'ideal' interventions. Korb et al. (2004) refer to them as 'independent' interventions.

4. This is an assumption we make here for the theorems that follow. It is not an assumption that is necessary in general. There may well be so called 'fat-hand' interventions that imply correlated interventions on two variables. But in such cases the discovery procedure has to be adapted and the following results do not hold in general.

5. We do not take 'exogenous' to be synonymous with 'uncaused' in general, but rather to represent a weaker notion. However, for brevity we will skip that discussion here.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

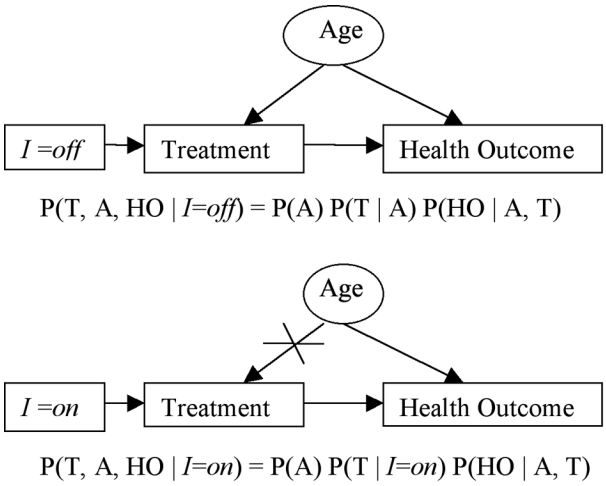


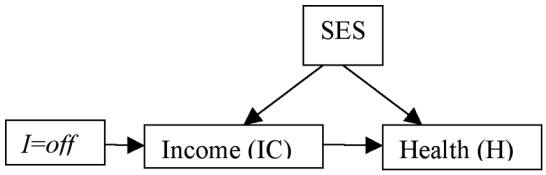
Figure 2.—Causal Bayes Net representing a 'structural' intervention (e.g., a medical trial) where Treatment (T) and Health Outcome (HO) are confounded by an unmeasured common cause Age (A). When the intervention variable *I* is set to *off*, the passive observational distribution obtains over the variables. When it is set to *on*, the structural intervention breaks any arrows incident on the treatment variable, thereby destroying any correlation due to unmeasured common causes.

The epistemological advantages of structural interventions on one variable are at least the following:

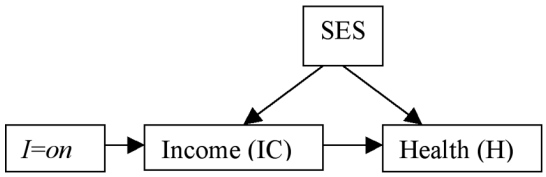
- No correlation between the manipulated variable and any other non-manipulated variable in the resulting distribution is due to an unmeasured common cause (confounder).
- The structural intervention provides an ordering that allows us to distinguish the direction of causation, that is, it distinguishes between $A \rightarrow B$ and $A \leftarrow B$.
- The structural intervention provides a fixed known distribution over the treatment variable that can be used for further statistical analysis, such as the estimation of a strength parameter of the causal link.

This is not the only type of intervention that is possible or informative, however. There may also be 'soft' interventions that do not remove edges, but simply modify the conditional probability distributions of the intervened upon variable. In a causal Bayes net, such an intervention would still be represented by an exogenous variable with a single arrow into the variable it intervenes on. Again, when it is set to *off*, the passive observational distribution obtains; but when it is set to *on*, the distribution of the variable conditional on its causes (graphical parents) is changed, but

https://doi.org/10.1086/525638 Published online by Cambridge University Press



$$P(IC,SES,H \mid I{=}off) = P(SES)\,P(IC|SES)\,P(H|SES,IC)$$



$$P(IC,SES,H \mid I{=}on) = P(SES)P(IC|SES, I{=}on)P(H|SES,IC)$$

Figure 3.—Causal Bayes Net representing a 'parametric' intervention. When the intervention variable I is set to *off*, the passive observational distribution obtains over the variables. When it is set to *on*, the parametric intervention does not break any arrows incident on the intervened variable, but only changes the conditional distribution of that variable. So, $P(IC|SES) \neq P(IC|SES, I = on)$. In contrast to structural interventions, parametric interventions do not break correlations due to unmeasured common causes.

their causal influence (the incoming arrows) are not broken. We refer to such an intervention as a 'parametric' intervention, since it only influences the parameterization of the conditional probability distribution of the intervened variables on its parents, while it still leaves the causal structure intact.[6] The conditional distribution of the variable still remains a function of the variable's causes (parents).

More formally, we have the following definition for a *parametric* intervention I_p on a variable X in a system of variables $\boldsymbol{V}$ (see Figure 3):

- I_p is a variable with two states (*on*/*off*).
- When I_p is *off*, the passive observational distribution over $\boldsymbol{V}$ obtains.
- I_p is a direct cause of X and only X.
- I_p is exogenous, that is, uncaused.
- When I_p is *on*, I_p does not make X independent of its causes in $\boldsymbol{V}$ (does not break the edges that are incident on X). In the factored joint distribution $P(\boldsymbol{V})$, the term $P(X|\text{parents}(X))$ is replaced with

6. As mentioned above, Korb et al. (2004) refer to such an intervention as a 'dependent' intervention. It has also been referred to as 'soft' (Campbell 2006) or 'conditional'.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

the term $P(X|\text{parents}(X), \quad I_p = on)$,[7] and otherwise all terms are unchanged.

There are several ways to instantiate such a parametric intervention. If the intervened variable is a linear (or additive) function of its parents, then the intervention could be an additional linear factor. For example, if the target is *income*, the intervention could be to boost the subject's existing income by \$10,000/year. In the case of binary variables, the situation is a little more complicated, since the parameterization over the other parents must be changed, but even here it is possible to perform a parametric intervention, for example, by inverting the conditional probabilities of the intervened variable when the parametric intervention is switched to *on*.[8]

3. Results. While a structural intervention is extremely useful to test for a direct causal link between two variables (this is the focus in the statistics literature), it is not straight forwardly the case that structural interventions on single variables provide an efficient strategy for discovering the causal structure among several variables. The advantage it provides, namely making the intervened upon variable independent of its other causes, is also its drawback. In general, we still want a theory of causal discovery that does not rely upon an a priori separation of the variables into treatment and effect as is assumed in statistics. Even time ordering does not always imply information about such a separation, since we might only have delayed measures of the causes.

Faced with a setting in which any variable may be a cause of any other variable, a structural intervention of the wrong variable might then not be informative about the true causal structure, since even the manipulated distribution could have been generated by several different causal structures.

For example, consider Figure 4. Suppose the true but unknown causal graph is (1). A structural intervention on C would make the pairs A-C and B-C independent, since the incoming arrows on C are broken in the postmanipulation graph (2). The problem is that the information about the causal influence of A and B on C is lost. Note also, that an association between A and B is detected but the direction of the causal influence cannot be determined (hence the representation by an undirected edge). The manipulated distribution could as well have been generated by graph (3), where the true causal graph has no causal links between A and C or B and C. Hence, structural interventions also create Markov equivalence

7. Obviously, $P(X \mid \text{parents}(X)) \neq P(X \mid \text{parents}(X), I_p = on)$.

8. We are grateful to Jiji Zhang for this example.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

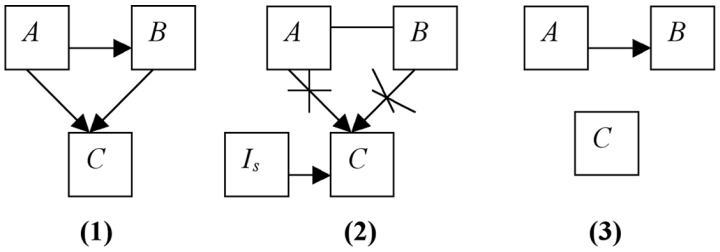


Figure 4.—Actual and manipulated causal graphs.

classes of graphs, that is, graphs that have a different causal structure, but imply the same conditional independence relations. (1) and (3) form part of an interventional Markov equivalence class under a structural intervention on C (they are not the only two graphs in that class, since the arrow between A and B could be reversed as well). Discovering the true causal structure using structural interventions on a single variable, and to be guaranteed to do so, requires a sequence of experiments to partition the space of graphs into Markov equivalence classes of unique graphs. Note that a further structural intervention on A in a second experiment would distinguish (1) from (3), since A and C would be correlated in (1) while they would be independent in (3).

Eberhardt, Glymour, and Scheines (2006) showed that for N causally sufficient variables $N - 1$ experiments are sufficient and in the worst case necessary to discover the causal structure among a causally sufficient set of N variables if at most one variable can be subjected to a structural intervention per experiment assuming faithfulness. If multiple variables can be randomized simultaneously and independently in one experiment, this bound can be reduced to $\log_2(N) + 1$ experiments (Eberhardt et al. 2005). These bounds both assume that an experiment specifies a subset of the variables under consideration that are subject to an intervention and that each experiment returns the independence relations true in the manipulated population, that is, issues of sample variability are not addressed.

Parametric interventions do not destroy any of the causal structure. However, if only a single parametric intervention is allowed, then there is no difference in the number of experiments between structural and parametric interventions:

> **Theorem 1**. $N - 1$ experiments are sufficient and in the worst case necessary to discover the causal relations among a causally sufficient set of N variables if only one variable can be subject to a *parametric* intervention per experiment. (Proof sketch in Appendix.)

https://doi.org/10.1086/525638 Published online by Cambridge University Press

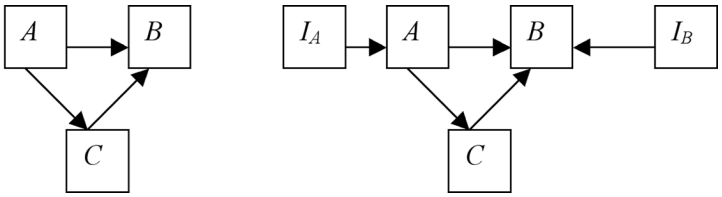


Figure 5.—A true unknown causal graph and a postmanipulation graph.

For experiments that can include simultaneous interventions on several variables, however, we can decrease the number of experiments from $\log_2(N) + 1$ to a *single* experiment when using parametric interventions:

> **Theorem 2**. One experiment is sufficient and (of course) necessary to discover the causal relations among a causally sufficient set of N variables if multiple variables can be simultaneously and independently subjected to a *parametric* intervention. (Proof sketch in Appendix).

The following example, illustrated in Figure 5, explains the result: The true unknown complete graph among the variables A, B, and C is shown on the left. In one experiment, the researcher performs simultaneously and independently a parametric intervention on A and B (I_A and I_B, respectively, shown on the right). Since the interventions do not break any edges, the graph on the right represents the postmanipulation graph. Note that A, B, and I_B form an unshielded collider,[9] as do C, B, and I_B. These can be identified (see footnote 9) and hence determine the edges and their directions A to B and C to B. The edge A to C can be determined since (i) A and C are dependent for all possible conditioning sets, but (ii) I_A, A and C do *not* form an unshielded collider. Hence we can conclude that (from [i]) there must be an edge between A and C and (from [ii]) that it must be directed away from A. We have thereby managed to discover the true causal graph in one experiment. Essentially, adjacencies can be determined from observational data alone. The parametric interventions set up a 'collider test' for each triple I_X, X, and Y with X—Y adjacent, which orients the X—Y adjacency.

4. Discussion. These results indicate that the advantage of parametric

9. Variables X, Y, and Z form an unshielded collider if X is a direct cause of Y, Z is a direct cause of Y and X is not a direct cause of Z and Z is not a direct cause of X. Unshielded colliders can be identified in statistical data, since X and Z are unconditionally independent, while they are dependent conditional on Y.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

TABLE 1. Parametric and Structural Interventions

Number of Experiments for Different Types of Interventions	Single Intervention per Experiment	Multiple Simultaneous Interventions per Experiment
Parametric interventions	$N - 1$	1
Structural interventions	$N - 1$	$\log_2(N) + 1$

interventions lies with the fact that they do not destroy any causal connections. (See Table 1.)

The theorems tempt the conclusion that parametric interventions are always better than structural interventions. But this would be a mistake since the theorems hide the cost of this procedure. First, determining the causal structure from parametric interventions requires more conditional independence tests with larger conditioning sets. This implies that more samples are needed to obtain a similar statistical power on the independence tests as in the structural intervention case. Second, the above theorems only hold in general for causally sufficient sets of variables. A key advantage of randomized trials (structural interventions) is their robustness against latent confounders (common causes). Parametric interventions are not robust in this way, since they do not make the intervened variable independent of its other causes. This implies that there are cases for which the causal structure cannot be uniquely identified by parametric interventions.

The two graphs in Figure 6 over the observed variables A, B and C with latent common causes L_1, L_2, and L_3 are indistinguishable given parametric interventions on A, B, and C. There is no conditional independence relation between the variables (including the intervention nodes omitted for clarity in the figure) that distinguishes the two graphs.[10] However, it is not always the case that causal insufficiency renders parametric interventions useless as the example in Figure 7 shows.

The parametric intervention I_X on X will not break the association between X and Y that is due to the unmeasured common cause L. This does not mean that the edge X to Y cannot be identified, since (a) if there were no edge between X and Y, then I_X and Y *would not* be associated, and (b) if the edge were from Y to X, then I_Y and X *would* be associated.[11]

10. In order to prove this claim, a large number of conditioning sets have to be checked, which we did using the Tetrad program (causality lab: http://www.phil.cmu.edu/projects/causality-lab/), but omit here for brevity.

11. Note, that it is not *simply* the creation of unshielded colliders that is doing the work here. If there were no parametric intervention on Y, we could not distinguish between $X \leftarrow Y$ and no edge between X and Y at all, *even though* the first case would create an unshielded collider $I_X \rightarrow X \leftarrow Y$. In the presence of latent variables, creating colliders is necessary, but not sufficient (as the more complex example shows).

https://doi.org/10.1086/525638 Published online by Cambridge University Press

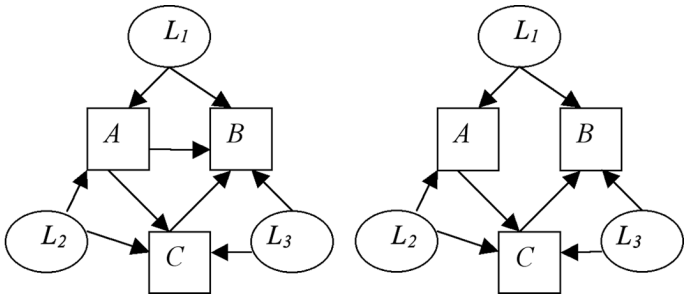


Figure 6.—Two graphs with latent variables that are indistinguishable by parametric interventions on A, B, and C (the intervention variables are omitted for clarity).

As the two examples show, parametric interventions are *sometimes* subject to failure when causal sufficiency cannot be assumed, but this depends very much on the complexity of the model.[12] It follows that the assumption of causal sufficiency does not compensate entirely for what parametric interventions lack in robustness to unmeasured common causes. First, the results for structural interventions also depend on causal sufficiency. (See Eberhardt et al. 2005.) A logarithmic bound on the number of experiments when using structural interventions would not be achievable in the worst case if there are latent common causes. Second, there are cases (as the example above shows) where parametric interventions are still sufficient to recover the causal structure even if there are latent common causes.[13] In some cases one can even identify the location of the latent variables. However, this is not always the case and we do not have any general theorem of these cases to report at this stage. What seems to play a much larger role than causal sufficiency is the requirement of exogeneity of the intervention. It results in additional independence constraints (creating

12. Note also the close similarity between parametric interventions and instrumental variables as they are used in economics. However, in economics—as in statistics—there is the general assumption that some kind of order is known between the variables, so that one only tests for presence of an adjacency, while its direction is determined by background knowledge.

13. It is not the simplicity of the second example (with just two variables) that allows for the identifiability of the structure when using parametric interventions in the presence of latent common causes. In fact, if we remove the edge $C \leftarrow B$ (instead of $A \leftarrow B$) in the three variable example above, the graph would be distinguishable from the complete graph among A, B, and C even if there are latent common causes for every pair of variables. However, it may not form a singleton Markov equivalence class under parametric interventions.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

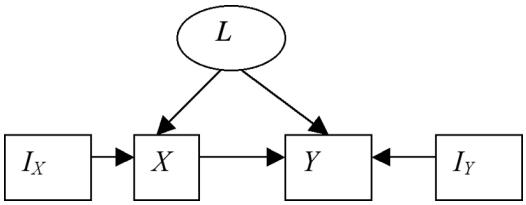


Figure 7.—Even in the presence of a latent variable, this causal structure can be uniquely determined with parametric interventions on X and Y in a single experiment.

unshielded colliders) whose presence or absence make associations identifiable with direct causal connections.

The moral is that while the types of interventions are not independent of the assumptions that are made about causal sufficiency, they are not interchangeable with them either. It is not the case that causal sufficiency and parametric interventions achieve the same search capability as causal insufficiency and structural interventions. The tradeoffs are rather subtle and we have only shown a few examples of the interplay of some of the assumptions. But there does not seem to be any general sense in which one can speak of 'weaker' or 'stronger' interventions with regard to their epistemological power.

Appendix: Intervention Variables

We represent interventions in causal Bayes nets by so called policy variables, which have two states. These policy variables are not causal variables in the sense of the other variables in the causal graph (although they may model instrumental variables): they need not have a marginal probability distribution over their values. By 'switching' the policy variable to one of its states we refer to the decision to perform an intervention. In the case of a parametric intervention, the policy variable creates unshielded colliders, while we claim that this is not the case for structural interventions. This will need some clarification, since this aspect is very relevant to the discovery procedures.

If our data contain samples both for when a variable X is subject to a structural intervention *and* when it is passively observed, then—using the sample distribution of I_s—we will find that even in the case of structural interventions, $I_s \rightarrow X \leftarrow Y$ forms an unshielded collider. The adjacency is obtained from the subsample where $I_s = off$, while the direction is determined when $I_s = on$. This sample constitutes a mixture of populations, one manipulated and one unmanipulated, as we may find it in a random-

https://doi.org/10.1086/525638 Published online by Cambridge University Press

ized trial with a control group that is not subject to any intervention. However, in studies where each condition in the randomized trial involves an intervention[14] (e.g., comparative medical studies) we do not have such a passively observed control group that would capture the unmanipulated structure. So the interventions are different (with regard to the manipulated distribution they impose) in different conditions, but if both conditions constitute *structural* interventions, then the causal structure cannot be recovered from the sample in all cases. Furthermore, if we perform multiple simultaneous structural interventions and only obtain data where either *all* intervention variables are *on* or all intervention variables are *off* for each sample, then again we cannot recover the causal structure in all cases.

In contrast, we can consider the unshielded collider in experimental designs involving *parametric* interventions since we can discover the unshielded collider even in a data set where the sample $I_p = on$ for all samples, as long as there are different '*on*'-states. And we can discover unshielded colliders in the case of multiple simultaneous parametric interventions, that is, when we only obtain data where either *all* intervention variables are *on* or all intervention variables are *off* for each sample. The key is that parametric interventions can be combined independently and performed simultaneously without interfering, while this is not the case for structural interventions.

Theorem 1 (proof sketch). Let each of the $N - 1$ experiments E_i, with $0 < i < N$ consist of a parametric intervention on X_i. In each case, the intervention variable I_i forms an unshielded collider with any cause of X_i. Hence, for any variable Y, where Y and I_i are unconditionally independent, but dependent conditional on C union $\{Y\}$ for all conditioning sets C, we know that Y is a cause of X_i. Further, in each experiment we check whether X_i and X_N (which is not subject to an intervention) are dependent for all conditioning sets. If so, then X_i is a cause of X_N. Since we perform a parametric intervention on $N - 1$ variables, all causes of these $N - 1$ variables can be discovered, and since we check for each variable whether it is a cause of X_N, all its causes are determined as well. Hence, $N - 1$ experiments are sufficient to discover the causal structure.

$N - 1$ experiments are in the worst case necessary, since $N - 2$ parametric interventions would imply that two variables are not subject to an intervention. This would make it (in the worst case) im-

14. In this case the '*on*'-state ($I_s = on$) of the intervention variable would have to be augmented to different *on*-states.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

possible to determine the direction of any edge between them, if there were one.

Theorem 2 (proof sketch). Since the parametric interventions described in the proof of the previous theorem do not interfere with each other, they can be performed all at the same time. That is, in a single experiment $N - 1$ variables would be subject to a parametric intervention and the causal structure could be discovered all in one go. One experiment is necessary, since—in the case of a complete graph—there would not be any unshielded colliders that would allow for the determination of the direction of any causal link between the variables. This is where the parametric $N - 1$ interventions are necessary. The theorem can also be derived simply from a theorem on rigid indistinguishability in Spirtes, Glymour, and Scheines (1993, Theorem 4.6).

REFERENCES

Campbell, J. (2007), "An Interventionist Approach to Causation in Psychology", in A. Gopnik and L. Schulz (eds.), *Causal Learning: Psychology and Computation*. Oxford: Oxford University Press.

Cartwright, Nancy (2001), "What is Wrong with Bayes Nets?", *The Monist* 16: 242–264.

——— (2002), "Against Modularity, the Causal Markov Condition and Any Link between the Two", *British Journal of Philosophy of Science* 53: 411–453.

——— (2006), "From Metaphysics to Method: Comments on Manipulability and the Causal Markov Condition", *British Journal of Philosophy of Science* 57: 197–218.

Eberhardt, Frederick, Clark Glymour, and Richard Scheines (2005), "On the Number of Experiments Sufficient and in the Worst Case Necessary to Identify All Causal Relations among N Variables", in Fahiem Bacchus and Tommi Jaakkola (eds.), *Proceedings of the 21st Conference on Uncertainty and Artificial Intelligence.* Corvallis, OR: AUAI Press, 178–184.

——— (2006), "N−1 Experiments Suffice to Determine the Causal Relations among N Variables", in Dawn E. Holmes and Lakhmi C. Jain (eds.), *Innovations in Machine Learning*, Theory and Applications Series: Studies in Fuzziness and Soft Computing. Vol. 194. New York: Springer-Verlag.

Fisher, Ronald A. (1935), *The Design of Experiments*. New York: Hafner.

Hausman, Dan, and James Woodward (1999), "Independence, Invariance, and the Causal Markov Condition", *British Journal for the Philosophy of Science* 50: 521–583.

——— (2004), "Modularity and the Causal Markov Condition: A Restatement", *British Journal for the Philosophy of Science* 55: 147–161.

Hitchcock, Chris (2007), "On the Importance of Causal Taxonomy", in Alison Gopnik and Laura Schulz (eds.), *Causal Learning: Psychology, Philosophy and Computation.* New York: Oxford University Press, 107–116.

Hoover, Kevin (2003), "Nonstationary Time Series, Cointegration, and the Principle of Common Cause", *British Journal of Philosophy of Science* 54: 527–551.

Korb, Kevin, Lucas Hope, Ann Nicholson, and Karl Axnick (2004), "Varieties of Causal Intervention", in *Proceedings of the Pacific Rim International Conference on AI.* New York: Springer.

Pearl, Judea (2000). *Causality.* Oxford: Oxford University Press.

Sober, Elliott (2001), "Venetian Sea Levels, British Bread Prices, and the Principle of Common Cause", *British Journal of Philosophy of Science* 52: 331–346.

https://doi.org/10.1086/525638 Published online by Cambridge University Press

Spirtes, Peter, Clark Glymour, and Richard Scheines (2000), *Causation, Prediction and Search*. Springer Lecture Notes in Statistics. Cambridge, MA: MIT Press.
Steel, Dan (2005), "Indeterminism and the Causal Markov Condition", *British Journal of Philosophy of Science* 56: 3–26.
Woodward, James (2003), *Making Things Happen*. Oxford: Oxford University Press.

https://doi.org/10.1086/525638 Published online by Cambridge University Press