

A SURVEY OF WOODWARD'S COUNTERFACTUAL THEORY OF CAUSATION¹

Ivan Camilo Verano
Universidad Nacional de Colombia
iverano@cmu.edu

Resumen: En su libro *Making Things Happen* (2003), James Woodward propone una teoría intervencionista de la causalidad. Esta teoría adopta como elemento esencial el análisis contrafáctico de aserciones causales. ¿Cómo debe entenderse e interpretarse un contrafáctico desde esta teoría intervencionista? Para responder esta pregunta, tomaré como guía la propuesta de David Lewis en su artículo "Causation" (1979). El objetivo de este artículo es comparar y diferenciar las propuestas de Lewis y Woodward; se mostrará como Woodward logra evitar algunas objeciones estándar a la teoría de Lewis, así como algunos nuevos problemas que Woodward debe enfrentar. El artículo se divide en tres partes. Primero, presentaré brevemente las ideas centrales de la teoría intervencionista de la causalidad; en seguida, explicaré la necesidad de adoptar un análisis contrafáctico en dicha teoría, esbozando de pasada la teoría contrafáctica de la causalidad de Lewis y sus diferencias con Woodward. Finalmente, en la tercera parte examinaré la noción de *intervención*, el concepto central en la teoría intervencionista, desde la perspectiva contrafáctica, para así presentar un problema fundamental en la teoría de Woodward.

Palabras clave: causalidad, intervención, contrafácticos, Woodward, Lewis.

Abstract: In his book *Making Things Happen* (2003), James Woodward proposes an interventionist account of causation. Such account requires the adoption of a counterfactual analysis of causal claims. However, how should counterfactual claims be understood and interpreted from an interventionist standpoint? I will try to answer this question taking as a guide the influential account of counterfactuals presented by David Lewis in his famous paper "Causation" (1979). The aim of this paper is to outline some general considerations that show how Woodward's counterfactual analysis differs from Lewis's, thus gaining immunity against classical objections to Lewis's analysis, and to present some difficulties of Woodward's own approach. The paper is divided in three parts. First, I will present briefly the main ideas of the manipulability theory of causation; in the second part, I will introduce the necessity of adopting a counterfactual analysis in such an account, followed by a brief presentation of Lewis's theory as well as the differences and similarities with Woodward's approach. Finally, in the third part I will examine the notion of intervention, a key concept in the manipulability account, from the standpoint of counterfactual analysis, and present a problem with Woodward's overall account.

Keywords: causation, manipulation, counterfactuals, Woodward, Lewis

¹ This paper is the result of a seminar on causation at Universidad de Los Andes under the direction of professor Andrés Páez, and a revised version was presented at the *First Colombian Conference in Logic, Epistemology and Philosophy of Science* on November 2009. I would like to thank professor Páez for his valuable criticism and patience.

I. INTERVENTION AND CAUSATION

The main idea behind the manipulability theory of causation can be summarized in the following principle: X is a cause of Y if manipulation on X brings about manipulation on Y. More precisely, X causes Y if intervention on X's *values* constitutes an intervention on Y's values. As an example, take the intervention on the possible values of a light switch (up-down) as bringing about intervention on a light bulb's possible values (on-off).² While this appears straightforwardly as a circular account of causation, explaining the notion of "cause" by means of the concept of "intervention",³ Woodward clarifies that his account is not meant to be reductive, but illuminating. That is, one of his main goals is not to reduce the concept of causation to another, more basic or well-defined non-causal concept, but rather to elucidate the content of causal assertions in scientific and everyday contexts. Thus, according to Woodward, the content of causal utterances is not so much about transference of energy or similarity between possible worlds,⁴ but rather about information susceptible to possible manipulation. We will not consider the extensive defense of this assertion, but rather its consequences. It is important to note, however, that Woodward's project is not a mere description of standard usage of the concept of causation; it has a normative, revisionary aspect to it that stipulates what people *ought* to mean when uttering causal claims.

Such clarification of what is meant by the concept of "causation" is called for in the general context of explanation. According to Woodward, what is characteristic of a causal explanation is that it shows how an event, the *explanandum*, depends on a set of different factors, the *explanans*. The goal of causal explanations is thus to provide information susceptible of manipulation; there is a practical side to the genesis of the concept of "cause". Woodward, in a naturalistic trend, presents examples taken from empirical psychology to conclude that the concept of causation was probably formed due to its practical benefits for survival. Such practical benefit is evidenced in the fact that because of our possession of the relation of causation, we can manipulate or exert control over one of the *relata* (effects) via the other *relata* (causes). As Dickenson & Shanks (1995) put it, it is "the capacity to control rather than just react to the environment that provided the impetus for the evolution of a mind and a nervous system capable of representing causality" (quoted in Woodward 2003 34). This remark seeks to place humans into a Skinnerian conditioned environment rather than in the classical Pavlovian conditioned environment, and emphasizes the importance of action and control in the content of the concept of causation over the simple association of a passive, empiricist approach: the Humean condition is not necessarily the human condition after all.

Without further exposition of the technical details, let us now list the central conditions and definitions of the manipulability theory of causation:

I. *Sufficient Condition*: If there is a possible intervention that changes X's values in a way such that performing this intervention (and no other) changes Y's values, then X causes Y. [sc]

² Woodward also admits non-binary values as mappings onto the real numbers, a mapping that doesn't need to be that of a mathematical *function* (it's not one-to-one, i.e. different real numbers can yield the same value).

³ To intervene can be understood as "to break the causal chain". Therefore, "intervention" is a causal concept itself, motivating the circularity argument exposed.

⁴ This refers to two reductive theories of causation: Salmon's 'mechanical' theory and Lewis's counterfactual theory.



A Survey of Woodward's Counterfactual Theory of Causation

2. *Necessary Condition*: If X causes Y, then there is a possible intervention on X's values such that performing it changes Y's values. [NC]
3. *Total Cause*: X is a total cause of Y if, and only if, there is a possible intervention on X that changes the values of Y. [DC]
4. *Direct Cause*: X is a direct cause of Y if there is a possible intervention on X that changes Y's values, while the remaining events Z_n of the causal system are fixed on a specific value by means of interventions (*ceteris paribus* clause). [DC]
5. *Contributive Cause*: If X is a contributive cause of Y, then either (a) X is a direct cause of Y, or (b) there is a causal chain $X-Z_1-Z_2-\dots-Z_n-Y$ such that between each successive link there is a direct cause relation. [CC]

Now, given this general outline of a manipulability theory of causation, we must ask a key question that any manipulability approach has to face: What happens when intervention is, for practical, technological, or moral reasons, not possible? Given the above conditions, it would appear that given such a case where intervention is not within our reach, no causation is present, since *ex hypothesi* there would be no information susceptible of manipulation. This conclusion would directly contravene the goal of the conceptual elucidation that Woodward advocates, because we would be forced to deny, for example, that there is a causal link between the Moon and the rise of the tides. A major *desideratum* of Woodward's account is to avoid contradiction, whenever possible, with common-sense, scientific and everyday causal assertions, given that his goal is semantic elucidation. In addition, the manipulability account seems to imply a sort of anthropomorphism insofar as it relies on a notion of intervention that refers to *our* interventions, or interventions performed by *us*, the human species.⁵ To solve these issues, two tactics will be necessary: on the one hand, we will need to introduce counterfactual analysis; on the other hand, a precise formal definition of the notion of "intervention" is called for. In the next section we will deal first with the introduction of counterfactual analysis.



2. COUNTERFACTUALS AND CAUSATION

It is uncontroversial, however, that causal relationships exist and that explanation is possible in circumstances in which actual manipulation is impossible, whether for practical or other sorts of reasons (...) The notion of information that is relevant to manipulation thus need to be understood modally or counterfactually. (Woodward 2003 10)

We have seen the necessity of adopting counterfactual claims in order to comply with the goal of semantic elucidation. So, to retake the example of the last section, we would have that the Moon actually is the cause of the rise of the tides, even if we are currently unable to intervene in its orbital path or mass quantity, because *if we were* to intervene on its orbit or constitution, the tides *would* change. The sort of counterfactual claims that can help us elucidate the semantic content of causal claims are, therefore, counterfactual claims *systematically associated* with *hypothetical* interventions (Woodward 2003 122). The association, as we will see eventually, is heavily based on abstract models, such as structural equations and graphs. Let us now recast the conditions of the manipulability account in terms of counterfactual claims:

⁵ I will not treat this objection here. See Woodward's way out on (2003 98).

1. *Sufficient Condition*: If there is a possible intervention that changes X's values such that, if it were to be carried out, it would change the values of Y, then X causes Y. [SC]
2. *Necessary Condition*: If X causes Y, then there is a possible intervention that changes X's values such that, if it were to be carried out, it would change Y's values. [NC]

A similar translation into counterfactual form can be made on the definitions of the different types of causes. In doing so, we can see that the manipulability account is, essentially, a counterfactual account of causation, insofar as every true causal claim is systematically associated with a true counterfactual claim. It is for this reason that Woodward states that a satisfactory causal explanation ought to answer a *what-if-things-had-been-different question* by providing a true counterfactual claim where the *explanandum* varies in accordance to variation in the *explanans* (2003 II). Let us now recast the relationship between a causal relation and an associated true counterfactual claim using the standard notation in the analysis of counterfactuals:

1. sc^* : If " $X \Box \rightarrow Y$ " is true, then X causes Y.
2. nc^* : If X causes Y, then " $X \Box \rightarrow Y$ " is true.

Here, the proposition " $X \Box \rightarrow Y$ " stands for the counterfactual relationship between Y's values and changes in X's values. This rephrasing of *sc* and *nc* constitutes a clear enunciation of the association between a causal claim and a counterfactual statement. It is clear, however, that not just any type of counterfactual will do. It is simply not true that the Moon causes the tides if the counterfactual "If I were to jog, I would be in a better shape" turns out to be true. The type of counterfactual claim associated to a causal claim must be *systematically* associated; Woodward phrases this as information about a *pattern* of counterfactual dependence (2003 11). What provides such information, or put in other words, what enables such systematic associations? The answer lies in a certain *appropriate understanding* of counterfactual claims (*ibidem*).

So far, we have shown that the manipulability theory of causation, *qua* counterfactual account, rests on a certain way of understanding and analyzing counterfactual claims, in a way such that the required systematic association between these sort of claims and causal claims is provided. So it is now natural to ask: How counterfactuals should be understood? What is this appropriate understanding of counterfactuals that is necessary to yield non-trivial, non-spurious causal relations? What are the truth-conditions of counterfactual claims appropriately understood? These are not novel questions; 20th century metaphysicians have tried relentlessly to clarify the logic of these subjunctive claims. Let's briefly present the difficulty of interpreting counterfactual claims given the traditional logical apparatus.

The 'material implication', as defined by Whitehead & Russell (?), fails to properly capture the sense of counterfactual claims. The truth-conditions of a $P \supset Q$ proposition state that it is true whenever the antecedent P is false. However, not every subjunctive conditional is true given that the antecedent is false. For example, if X is such that it's impossible for it to ever happen, and Y did happen, then it's not obvious that the counterfactual "If X had happened, Y had happened" is true. C. I. Lewis's 'strict implication' also fails to represent what is meant by a



A Survey of Woodward's Counterfactual Theory of Causation

counterfactual claim in the context of causation. The reason is that we understand causation as a relation that holds in virtue of empirical happenings, not in virtue of logical or metaphysical reasons. That a proposition stating Y 's occurrence is *necessarily* implied by a proposition stating X 's occurrence, while could be capturing the concept of a necessary condition for causation between X and Y , fails to explain cases of sufficient causation. For example, I throw a bottle at a rock and it breaks; it is simply not true that the bottle's breaking was necessarily caused by my throwing it into a rock; it could have been also caused by my stomping on it. My throwing was sufficient, alas not necessary, for the bottle's breaking.

So the question remains: What are the truth-conditions of counterfactual claims, such as "If X had happened, Y had happened"? The search for criteria seems to be doomed to speculation, given not only the dependence of counterfactuals on context, but also the occasional vagueness of their meaning. Consider an example due to Quine: If Julius Caesar had been in charge of the UN's army in the Korean conflict, then he would have used (a) nuclear weapons, or (b) catapults. Is the counterfactual with (a) as a consequent true, and false otherwise? What must be accounted for to decide this? Julius Caesar's psychology? Perhaps his preference for novel weapons? Does it even make sense to imagine such a situation? Instead of trying to provide answers to these questions, let me focus on cases where it does seem to make sense to imagine counterfactual situations, and where the truth or falsity of counterfactual claims seems to be clearly stipulated, somewhere in the middle ground between material and strict implication (Collins *et al.* 2004: 3). David Lewis's influential approach will serve us as a guide towards clarifying Woodward's approach to counterfactual analysis.

Lewis's central notion in his analysis of counterfactuals, employing possible world semantics, is the notion of comparative similarity. It is defined as a relation R between possible worlds w within a set W of all the possible worlds, thus defining a triple $\langle W, R, w \rangle$. Comparative similarity then *orders*, albeit weakly,⁶ all the w_i with respect to the actual world. A world w_1 will be closer to the actual world than a world w_2 if it is more similar to the actual world than w_2 is. Similarity between worlds includes not only similarity with respect to general natural laws, but also with respect to concrete events. Given this modal approach, " $A \square \rightarrow B$ " is defined to be true if: (a) There are no possible worlds where A is true, or (b) a world w_i where A and B are true is closer to the actual world than any other world where A is true and B is false (1979: 163-166). Lewis then makes propositions A and B represent the occurrence of actual events, call them f and g respectively, in the standard form of $O(f)$ (" f occurs") and $O(g)$ (" g occurs"). Now Lewis is in a position to state that the event g depends counterfactually on the event f if, and only if, the counterfactual " $\neg O(f) \square \rightarrow \neg O(g)$ " is true. Now, since counterfactual dependence implies causation, we would then have that f would be the cause of g .

Independently of objections posed to Lewis's counterfactual analysis, let us now pick up Woodward where we left him and compare the compatibility of Lewis's account of counterfactuals with the manipulability account of causation. *Prima facie*, it's fair to say that Lewis's analysis is incompatible with Woodward's approach for two reasons: on the one hand, if we accept Lewis's analysis as an appropriate interpretation of counterfactual claims in the manipulability account,

⁶ This ordering is weak because there can be more than one possible world standing in the same degree of similarity with the actual world.



we would be inconsistent with the *semantic elucidation* goal that Woodward has set himself to achieve. It's quite implausible that scientific as well as everyday uses of counterfactual claims involve using Lewis's similarity metric (Woodward 2003 137). On the other hand, Lewis's approach is inevitably *reductive*. In a comparison with Lewis, Woodward states that "[Lewis's] idea is to define the notion of causation in terms of a more general notion of counterfactual dependence that does not itself presuppose causal notions" (2003 136). The manipulability account of causation cannot aspire to such a reduction, as its main tenets involve a notion in the conceptual vicinity of the concept of causation itself, *i.e.* the notion of *intervention*.



(Fig. 1)

Despite the crucial differences between Lewis and Woodward, there are nonetheless important points of similarity. One of these is the *verdict* that both Woodward's and Lewis's approaches would provide in regards to

the truth of counterfactual claims. To illustrate this point, consider a cause C that is the common cause of two effects, E₁ and E₂.

Let us now examine the truth value of the claim "If E₁ had not occurred, E₂ had not occurred", represented as " $\neg O(E_1) \textcircled{R} \neg O(E_2)$ ". What is at stake here is the problem of spurious correlations that yield false causal claims; while it may be true that E₁ and E₂ are correlated, it's false to state that one of them causes the other. Theories of causation must prove their worth by avoiding this problem, ruling the counterfactual claim above as false, and Woodward's and Lewis's theories are no exceptions. From the interventionist point of view, it's easy to appreciate the falsity of such counterfactual claim. It suffices to intervene and 'break the arrow'⁷ (*cf.* fig. 1) between C and E₁ while leaving the arrow from C to E₂ intact. This would show that there's no causation between the E's, as there's no 'causal arrow' between them given that intervening on one of them does not affect the other. From Lewis's perspective, we have that the possible world that is the closest to the actual world⁸ is one where C occurs and everything is happening just like it does in the actual world up to a point in time prior to E₁'s occurrence, when suddenly a miracle occurs and E₁ does not occur, but E₂ does. Now, since in this possible world the antecedent " $\neg O(E_1)$ " is true, but its consequent " $\neg O(E_2)$ " false, and since this is the closest world to the actual world, we have that the claim " $\neg O(E_1) \textcircled{R} \neg O(E_2)$ " is false, thus saving the problem of spurious correlations.

This example shows, and Woodward notes this, that Lewis's small, local miracles serve a similar purpose to that of interventions in the manipulability theory. It also shows how divergent both approaches are; Woodward has no use whatsoever of possible-world semantics, let alone a relation of similarity between worlds. This, in turn, stresses the question of what constitutes an *intervention*, and how are they characterized within the manipulability theory framework. This is the task of the next section.

⁷ Woodward employs directed graphs to represent relations of direct causation. Note that this 'breaking' intervention need not consist in an intervention on C's values, but some other exogenous node can be employed, e.g. a node D such that it causes E₁ to not-occur.

⁸ There is extensive literature on whether such a possible world is really the closest to the actual world. I will not dwell on this issue here.



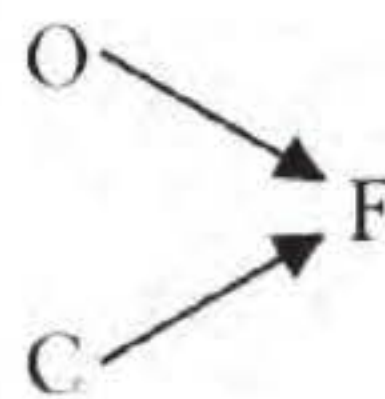
2. INTERVENTION AND COUNTERFACTUALS

Let us recall the counterfactual formulation we did above of the sufficient and necessary conditions for causation according to the manipulability account.

1. sc^* : If " $X \square \rightarrow Y$ " is true, then X causes Y .
2. nc^* : If X causes Y , then " $X \square \rightarrow Y$ " is true.

Let us now focus on what I consider to be the most interesting case, sc^* , where we go from a true counterfactual to a true causal claim. We have already seen that the types of counterfactuals involved are those associated with hypothetical interventions. Woodward suggests, as means of testing the truth-value of a counterfactual claim, to actually carry out the interventions that the antecedent of the subjunctive claim states. This intervention would examine if a correlation between the so-called cause and the so-called effect does hold (*i.e.* if information about a pattern is found). If it does hold in a significant number of cases, it would be fair to take the counterfactual being analyzed as true: "The existence of a correlation between X and Y that persists under the interventions specified in the antecedent of this counterfactual is in turn evidence that the counterfactual is true" (Woodward 2003 105). To *intervene* would be, therefore, to elucidate the content of a counterfactual claim confronted to causal claims. When one intervenes, the intervention contained in the antecedent of the subjunctive conditional is then actually carried out; if the consequent does not obtain, we would have evidence of its falsehood; if the consequent does obtain (under systematic interventions), we would have evidence of its truth. This test, then, is a sort of transformation of the counterfactual claim into a material implication, where our intervention sets the value of the antecedent in "true" and stipulates the truth-value of the whole claim in function of the consequent's truth-value.

We seem to be begging the original question, however, for how can such testing be performed in cases where *actual* intervention is not within our reach? These are the sort of cases that the manipulability account needs to save if it is to be of any interest. Woodward answers this question appealing to a different way of representing causal relationships other than directed graphs: systems of structural equations (often employed in physics as well as econometrics). Unlike directed graphs, these systems of equations not only point to the existence of causal relations among variables, but also encode information about how a variable would change as a function of change in some other variable (Woodward 2003 43). This resembles the notion of mathematical function, where given certain input, an output is given as a result of certain computations carried out following specific rules. This resemblance is best noted by Judea Pearl when in a footnote he tells us: "Every mathematical function is interpreted hypothetically, and the study of counterfactuals is merely a study of standard mathematical functions" (2000). Let us illustrate this point with the following example due to Woodward himself. Let's say we have three variables: O ("there is oxygen present"), S ("there is a short-circuit") and F ("there is fire"). Assume that the circumstances are such that both O and S are necessary causes of F , but separately not sufficient. A directed graph representation would be as fig. 2.



(Fig. 2)



The graph (*cf.* fig. 2), however, doesn't capture the fact that both O and C must happen in order for F to happen. As a matter of fact, the graph shows both O and C to be sufficient causes, which is not the case. A structural equation does a better job at representing such relations. If we assign the values of 0 and 1 to the variables according to whether there's oxygen or not, etc., we can see that the following equation captures the appropriate relations:

$$F = S \times O$$

Here we can see that in the absence of one of the necessary causes, the value of F will be zero regardless of the value of the other variable. In an analogous manner, if both causes do happen, then the result will be that the effect also happens (1×1 equals 1). A similar situation occurs in the case of equations in Physics. Even it is true that we are unable to actually intervene on the orbital path of the Moon, or on its mass, we can nevertheless perform a hypothetical intervention employing the mathematical models of Newtonian equations, such as the Law of Universal Gravitation:

$$F = G (M m/r^2)$$

Woodward would hold that this equation provides means to evaluate the truth of the counterfactual claim: "If the Moon were at twice the distance from Earth, it would exert less force on the tides". One would only have to employ the above equation and multiply r times two. Once the values of the variables are crunched in, and the resulting value of F is computed, all that is left is comparing the result of this F' with the 'normal' value of F (*i.e.* when r is not multiplied times two). If $F' < F$, then we have evidence of the truth of the counterfactual claim above. The fact that structural equations of this kind encode counterfactual information allows us to *see* what would happen if change in one or another variable were to take place, and in virtue of this is that we can judge the veracity or plausibility of counterfactual claims. So it is clear that the kind of interventions carried out when *actual* intervention is impossible is interventions in a *model*; given such a *model* of the actual world, we can introduce values on variables that actually were not assigned to them (our introduction would therefore be 'against the facts', *i.e.* counterfactual) and examine what result our model provides, whether it conforms or not to the consequent of a subjunctive conditional claim. The plausibility of this counterfactual claim, I think, is then derived from the merits of the model employed. Naturally, if it is an accepted, accurate model, then its results would give wider plausibility to claims verified by it. In a defense of manipulationist theories against Nancy Cartwright's (2002) objection that manipulationist accounts are "operationalist" (singling out just one procedure of testing), Woodward (2008) maintains that manipulationist accounts do not deny that some other procedures for testing causal claims are possible. Given the strict relation between causal claims and counterfactual claims, this seems to imply that there may be other procedures for testing counterfactual claims, compatible with the manipulationist approach. But what such an approach could be is not pursued by Woodward. This puts an enormous weight on the role played by representations of causal relations, *i.e.* directed graphs and more specially, structural equations.

However, it is interesting to note that we cannot guarantee the truth-value of a counterfactual being supported by these mathematical models, insofar as there is no way to tell if such models are



A Survey of Woodward's Counterfactual Theory of Causation

the models of how the actual world works. The truth-value assigned to the counterfactual claim would be just like the truth-value of a scientific hypothesis that has been verified a given number of times. Also, at this point, we can turn Woodward's purported goal with the manipulability theory against him. We have seen that such a theory relies on use of models, scientific ones presumably, to give an account of cases where actual manipulation is outside the scope of our capabilities. But is this really the content of causal claims? It may be obvious to be the content of causal claims within scientific practice, but is it also the use of models constitutive of the content of everyday causal claims? Woodward argued that Lewis's account wasn't in part appropriate for this reason; now his account could be said to also be insufficient on the same ground. It is simply not obvious, or even plausible, that one has to *know* such models to *know* that a given counterfactual is true or false. I can be sure that if I were to jump off a cliff, I would die, regardless of my knowledge of aerodynamics, Earth's gravitational constant or my body's biology. How does Woodward's account of finding out a counterfactual's truth constitute also an elucidation of the content of *all* causal claims needs further argumentation. But even if such argumentation is provided, it is hard not to think that it will be somewhat farfetched. Perhaps, it could be said, the normative aspect of Woodward's account enters here; when a subject utters a causal claim, she should be in a position to determine how the effect would change or vary under intervention on the cause; *i.e.* she should possess *information* about how such change remains invariant under some possible manipulations. But what would constitute this information? Possession of a structural model? Furthermore, if the manipulability theory's most interesting application, counterfactual cases, is inconsistent with the theory's main goal, semantic elucidation, it is not clear what value such theory can end up having. The fact that it is an essentially non-reductive account of causation would strip it of any philosophical interest, if the previous assertions turn out to be true.

REFERENCES

CARTWRIGHT, N.

(2002) "Against Modularity, the Causal Markov Condition, and Any Link Between the Two: Comments on Hausman and Woodward", *British Journal for the Philosophy of Science*, 53. 411-453.

COLLINS, J., N. HALL AND L. A. PAUL

(2004) "Counterfactuals and Causation: History, Problems and Prospects". In *Causation and Counterfactuals* J. Collins, N. Hall, & L. A. Paul (eds.). Cambridge: MIT Press. 1-58.

DICKINSON, A. AND D. SHANKS

(1995) "Instrumental Action and Causal Representation". In *Causal Cognition*. D. Sperber, D. Premack y A. Premack (eds.). New York: Oxford University Press. 5-25.

LEWIS, D.

(1986) "Causation". In *Philosophical Papers Vol. II*. New York: Oxford University Press. 159-171.
(1979).

MENZIES, P.

(2008) "Counterfactual Theories of Causation", *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2008/entries/causation-counterfactual/>>.



PEARL, J.

(2000) "The Logic of Counterfactuals in Causal Inference" In *Journal of American Statistical Association* (with discussions by G. Casella, D. Cox, S. Greenland, J. Pearl, J. Robins, D. Rubin, S. Schwartz, G. Shafer, and L. Wasserman), Vol. 95, No. 450, 428-435, June 2000.

WOODWARD, J.

(2003) *Making Things Happen*. Oxford: Oxford University Press.

(2008) "Causation and Manipulability", *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition), Edward N. Zalta (ed.). URL = <<http://plato.stanford.edu/archives/win2008/entries/causation-mani/>>.

