CHAPTER 5

Causal Pluralism and Force Dynamics

PHILLIP WOLFF

5.1 Overview

Causal relations permeate human knowledge across a wide range of domains, from the physical to the abstract. The existence of causal relations in many different domains raises an interesting challenge for accounts of causation. On the one hand, the word *cause*, and related words, can be used to describe a very wide range of situations, consistent with the possibility that people's representation of causation is based on a single notion of causation, a perspective often called *causal monism*. On the other hand, as discussed in this chapter, it has proven difficult to develop a unitary theory of causation capable of addressing the various phenomena associated with causation. The difficulty in developing such a theory points to the possibility that people represent causation in several different ways, a view commonly referred to as *causal pluralism* (Cartwright 2004, DeVreese 2006, Godfrey-Smith 2010, Longworth 2010, Psillos 2008, Reiss 2009, Lombrozo 2010, Hall 2004, Hitchcock 2003, Lakoff & Johnson 1999).

In this chapter I describe the main arguments in support of both approaches to the representation of causation, and ultimately conclude in favor of causal monism. The case for causal monism begins with the observation that causal pluralism has a major flaw. In particular, when the full range of phenomena associated with causation is examined, it becomes clear that these phenomena cannot be successfully addressed by simply combining different theories of causation because some of these phenomena remain out of reach of all theories of causation.

A second argument in favor of causal monism is that, in contrast to recent claims, a monistic approach remains a viable position once the notion of causation is understood in terms of a general approach to causation known as force dynamics (Talmy, 1988). As discussed below, force dynamics is able to account for a much wider range of phenomena than can be accounted for by other monistic theories. In particular, force dynamics can explain causation by omission, double prevention, transitive reasoning, and correspondence relationships between various causal expressions. Part of this discussion will involve a description of a particular version of force dynamics, called the *force theory* (Wolff 2007, Wolff, Barbey & Hausknecht 2010), and the way this theory addresses the various phenomena listed above. One last argument in favor of causal monism is the uniformity we see in how people construct causal relations from causal chains, a type of reasoning known as *causal composition*. In particular, I will discuss the results of several recent experiments showing that causal reasoning in the case of abstract causal relations mirrors that of causal reasoning in the case of concrete causal relations. The finding that causal composition for abstract and concrete relations occurs similarly is consistent with a unified notion of causation and not easily explained by pluralistic accounts of causation. Together, these various lines of evidence show not only how a monistic account can be successful, but also why a pluralist account may not be needed.

5.2 Pluralist approaches to causation

Causal pluralism has been rendered in several different ways. According to some theorists, causal understanding can be reduced to just two major kinds of causation (Hall, 2004). According to others, causal pluralism emerges from the existence of a wide range of specific kinds of causation (Cartwright, 2004). Still others view causal understanding as a prototype notion that admits to any number causal notions, provided they are sufficiently similar to the prototype (Lakoff & Johnson 1999, see also Psillos, 2008). The bulk of this chapter will focus on the version of causal pluralism that holds that causal understanding is based on two main types of causation, but towards the end of the chapter, other versions of causal pluralism will be considered.

In Hall's (2004) version of causal pluralism, causal understanding is based on both dependency and production causation. That is, according to Hall (2004), people's causal understanding is sometimes based on notions of dependency and other times upon notions production. The basic intuition behind dependency causation is that a cause is something that makes a difference, specifically, a difference from what would have happened had the cause not been present (Lewis, 1973). The intuition can be described in the form of a counterfactual dependency: event A causes B provided that if A had not occurred then B would not have occurred in the closest possible world to our own (Lewis 1973, Shibatani 1976). The closest possible world proviso holds when it is the case that 'if A then B' is more similar to the actual world than 'if A then ¬B.' The closest possible world condition is included in counterfactual accounts of causation in order to distinguish genuine from spurious causation and to determine causal direction (see Dowty, 1979).

Another type of dependency relationship is a statistical dependency relation. In terms of a statistical dependency, a cause is something that changes the probability of an effect (Hitchcock, 2010). A statistical dependency approach to causation does not require that an effect always follow from an occurrence of a cause, or that an effect only occur in the presence of one particular cause. Rather, it requires that occurrence of a cause makes a *difference* in the probability that the effect will occur. In the case of facilitative causation, it can be said that a cause *raises* the probability of an effect (Suppes 1970, Ellis 1991). A particular version of probability raising is instantiated in Cheng & Novick's (1992) probabilistic contrast model. In their account, a statistical dependency is present when the probability of an effect in the presence of a candidate cause, $P(E|C)$, is greater than the probability of the effect in its absence, $P(E|\neg C)$, that is, $P(E|C) > P(E|\neg C)$.

Dependency causation can be contrasted with production causation. The concept of production causation is based on the notion of transfer of quantities between the cause and effect (for a recent review, see Kistler, 2006a) as well as on the notions of causal power and force (for recent reviews, see Mumford, 2008, Wolff & Shepard, 2013). Theories that have emphasized the notion of transfer include Aronson's (1971) transference theory, which holds that causation is present when contact between two objects allows a quantity possessed by the cause (e.g., velocity, momentum, kinetic energy, heat, etc.) to be transferred to the effect. In Fair's transfer (1979) theory, causation occurs when physical quantities such as energy or momentum flow from the cause to the effect. In Dowe's (2000) conserved quantity theory, causal interactions involve a bidirectional exchange of energy. Kistler (2006a) develops a transfer theory of causation that, unlike the other theories, separates the concept of causation from time, hence allowing for a theory of events in terms of causation as well as for the representation of backwards causation.

The idea that causation might be based on causal power has recently been partially formalized in Mumford and Anjum's threshold account of causation (2011a, 2011b; see also Anjum & Mumford, 2010). In this theory, an effect is produced when the sum of various powers—represented as vectors—reaches a threshold for that effect. Another theory of causation that describes causation in terms of powers is Talmy's (1988) force dynamics theory, which describes such powers in terms of tendencies. As will be discussed in greater detail below, Talmy's theory has been extended and formalized by Wolff and his colleagues (Wolff 2007, Wolff, Barbey & Hausknecht 2010, Wolff & Zettergren 2002). In a related proposal by Fales (1990), the notion of causation is linked to the notion of force as understood through our bodily experiences with forces (see also White 2012, Wolff & Shepard 2013). Finally, there have been several recent accounts of the expression of causation in language that build off of notions of force and energy. In Gärdenfors and colleague's two-vector model, causal expressions are represented in terms of, at least, one result vector, one force vector, and a central object that undergoes a change (Warglien, Gärdenfors & Westera, 2012). In a model-theoretic theory developed by Copley & Harley (submitted; this volume), the semantics of causation are grounded in the notion of force, and specified formally as a function functions from an initial situation to a result situation, *ceteris paribus*. In the broader psychological and philosophical literatures, theories of causation based on causal powers or forces are sometimes referred to as *causal process* theories (Dowe 2007, Wolff, Barbey & Hausknecht 2010) or *mechanism* theories (Ahn & Kalish 2000, Craver 2007, Machamer, Darden & Craver 2000).

5.3 Problematic phenomena for dependency and production theories

Dependency and production causation have different properties. One of the key features of dependency causation is that the manner in which the effect is brought about is not explained. This property makes dependency causation well suited for handling particular types of causation, for instance, *causation by omission,* that is, causation that results from an absence (McGrath, 2003, 2005; Schaffer, 2000). Causation by omission is exemplified by statements such as 'Lack of caffeine caused a headache' or 'The absence of water caused the plant to die.' Such causation has traditionally been difficult to explain in terms of process causation since, plainly, energy or force cannot be transmitted from an absence (Mumford & Anjum 2011a, Schaffer 2000, Wolff, Barbey & Hausknecht 2010).

Another phenomenon that dependency causation explains well is *causation via double prevention* (Collins 2000, Dowe 2001, Hall 2000, 2004). Consider, for example, the well-known scenario of the Dutch boy who is protecting a town from flooding by keeping his finger on a leak in a dam (McGrath, 2003). In this scenario, if the boy were to remove his finger, he would, in effect, bring about a double prevention: removing his finger would prevent the prevention of the leak, and ultimately, the flooding of the town. To describe this scenario, we could say ‘The boy caused the flooding of the town.’ Most production theories do not readily explain why the boy can be viewed as a cause because the boy does not transmit energy to either the dam or water, and hence these theories cannot explain why the boy should be interpreted as a cause (Kistler 2006a, Schaffer 2000, Wolff, Barbey & Hausknecht 2010). Such causation is readily handled by dependency theories because they only insist on a dependency between the cause and the effect with no regard for how the causation comes about.

While a dependency approach to causation is able to address certain causal phenomenon, it also faces several challenges. First, in certain situations, a dependency approach to causation predicts causal relations where none exist. For example, a dependency approach to causation predicts that birth causes death since it is the case that if a person had not been born, that person could not have died. To borrow an example from Kim (1973), dependency theories predict that Mondays cause Tuesdays since it is the case that if yesterday had not been Monday, then today could not be Tuesday. A second major problem for dependencies theories is that they sometimes fail to predict causal relations when causation appears to be present. Such failures occur in cases of so-called ‘causal preemption’ (Thomason, 2013), ‘overdetermination’ or ‘redundant causation.’ This problem was explored empirically by Walsh & Sloman (2005), who presented participants with a version of a scenario developed by Hall (2004):

> There is a bottle on the wall. Billy and Suzy are standing close by with stones and each one throws a stone at the bottle. Their throws are perfectly on target. Billy happens to throw first and his reaches the bottle before Suzy's. The bottle breaks.

Participants overwhelmingly indicated that Billy caused the break. This result is problematic for a dependency view of causation because the breaking of the bottle was not dependent on Billy: if Billy had not thrown the stone, the bottle would still have broken due to Suzy. Walsh & Sloman (2005) interpret this finding as supporting a process account of causation in that the breaking of the bottle did not depend on Billy, but there was a transmission of conserved quantities from Billy to the bottle that made Billy the cause of the event.

A third major problem for dependency theories emerges in trying to account for causal transitivity. Causal relations appear to have the property of being transitive: if a causes b and b causes c, then we can say a causes c (Lewis 1973, 2000, Hall 2000, Barbey & Wolff 2007, Hall 2004, Wolff et al. 2010, Sloman, Barbey & Hotaling 2009). If, however, causal relations are understood in terms of counterfactual dependencies, then we are faced with a problem because counterfactual dependencies are not transitive (Collins 2006, Hall 2000, Lewis 1973, 2000). As noted by Stalnaker (1968), we could say that 'If J. Edgar Hoover had been born a Russian, he would have been a communist', and we could say that 'If J. Edgar Hoover had been a communist, he would have been a traitor', but transitive reasoning would lead to the unreasonable conclusion that 'If J. Edgar Hoover had been born a Russian, he would have been a traitor.' As argued by Hall (2004), a production view of causation seems to handle causal transitivity better: if the first domino causes the middle domino to fall, and the middle domino causes the last domino to fall, there is a transmission of energy from the first to last domino, which may explain why it is acceptable to conclude that the first domino caused the last domino to fall.

In sum, there are a number of phenomena, such as causation by omission and double prevention, that are problematic for most production theories, but not dependency theories, and other phenomena such as causal redundancy and transitivity that are problematic for the dependency theories, but not production theories. Causal pluralism explains all of these phenomena by holding that causation is not a single, unitary concept but rather that the concept of causation must sometimes be understood in terms of dependency and other times in terms of production. Given its greater explanatory power, a causal pluralism account would seem to provide a better account of causation than either dependency or production theories. However, when the range of causal phenomena is widened, it becomes apparent that there remain phenomena that, in most cases, defy explanation by either of the main classes of causation, implying that a complete account of causation cannot be achieved by simply cobbling together different theories of causation.

5.4 Problematic phenomena for causal pluralism

One phenomenon that generally cannot be handled by either dependency or simple production models is the distinction between CAUSE relations versus ENABLE and ALLOW relations (Wolff 2007, Wolff & Song 2003, Cheng & Novick 1991). These two concepts are similar in meaning but not interchangeable, as illustrated by the sentences in 1 and 2.

1 a. A blackout caused Peter to turn on the flashlight.

b. A switch allowed Peter to turn on the flashlight.

2 a. #A blackout allowed Peter to turn on the flashlight.
b. #A switch caused Peter to turn on the flashlight.

The sentences in 1a and 1b are perfectly acceptable. However, if the verbs are switched, the resulting sentences (sentences 2a and 2b) sound odd, demonstrating that the concepts of CAUSE and ENABLE differ in meaning. Neither dependency nor (most) production theories offer an explanation for this distinction (Wolff 2007, Goldvarg & Johnson-Laird 2001, Wolff & Song 2003). If causation is understood in terms of counterfactuals, the distinction between CAUSE and ENABLE is lost because both CAUSE and ENABLE relations involve dependency. With respect to the examples in (1), a counterfactual criterion predicts that a switch should make a fine cause because if the switch had not been present, Peter would not have been able to turn on the flashlight. However, as shown in sentence 2b, a switch is not easily viewed as a causer. Most production approaches to causation fare no better. A blackout does not transmit any more energy to Peter than does a switch, but a blackout is better viewed as a cause than is a switch.

A second phenomenon that cannot be explained by either a dependency or production approach to causation is the fact that negation can sometimes give rise to synonymy relationships between different kinds of causal expressions. Recent research has shown that causation, broadly construed, consists of three main kinds of causal concepts: CAUSE, ALLOW, ENABLE, and PREVENT (Wolff & Song, 2003). When the cause or effect in one of these types of causal relations is negated, the resulting statement can sometimes be paraphrased in terms of another type of causal relation (Wolff, et al., 2010). For example, a statement such as *Salt causes the* *absence of ice* has roughly the same meaning as in *Salt prevents ice*. Similarly, *Lack of wind in the room allowed the dust to settle* seems to imply *Presence of wind in the room prevents the dust from settling*. Other relationships are not so clear. When we say *Aspirin prevents clotting,* does this imply that *Lack of aspirin causes clotting*? Similarly, if *Green tea prevents Alzheimer's*, is it true that *Lack of green tea causes Alzheimer's*? Interestingly, as discussed in detail in Wolff et al. (2010), simple dependency theories predict that the latter two pairs of statements should be as synonymous as the first two pairs of statements. Most process theories are unable to handle negation altogether and so make no predictions about the relationship between different kinds of causal statements. In sum, simple dependency and process theories are unable to explain the various ways in which causal expressions are related to each other through the process of negation.

The reason why simple dependency and production theories cannot handle the phenomena just described is because their representations are too coarse. To distinguish CAUSE from ALLOW or to explain asymmetries in synonymy, a finer level of detail is required. In the next section I describe how this finer level of detail can be provided in terms of force dynamics. Force dynamics can be viewed as a type of production theory of causation, but unlike other production theories, force dynamics incorporates distinctions that allow it to explain all of the phenomena discussed so far, including the distinction between CAUSE and ENABLE, causation by omission, transitivity, and synonymy relationships. A force dynamic approach to causation demonstrates that a pluralist approach to causation need not be adopted in order to account for the range of phenomena associated with causal relations.

5.5 A force dynamics approach to causation

In prior work I have shown that causal relations can be understood in terms of configurations of forces (Wolff & Song 2003, Wolff 2007, Wolff et al. 2010). I refer to this account, which is based on Talmy's (1988) theory of force dynamics, as the *force theory*. One part of the theory specifies how people represent individual causal interactions in terms of configurations of force. A second part describes how people combine these configurations of force into chains. It is by means of combining configurations that the force theory is able to account for problematic phenomena described above.

5.5.1 Representing individual causal relations

According to the force theory, individual interactions involve two main entities: an affector and a patient (the entity acted on by the affector). The theory holds that people specify causal relations in terms of configurations of forces that are evaluated with respect to an endstate. An endstate can be conceptualized as a location in a state space that the patient might move towards or reach. The forces may be physical, psychological (e.g., intentions), or social (e.g., peer pressure) (Copley & Harley submitted, this volume, Wolff 2007). It is assumed that people's representations of forces are relatively accurate with respect to direction, but not magnitude. Uncertainty about the magnitudes of forces builds in a certain level of indeterminacy into people's representations of causation, which, in turn, can have consequences for how people describe chains of configurations of force involving multiple causal relations. This effect will be described in more detail below. For individual configurations of forces, uncertainty about the magnitudes of the forces is predicted to have no major consequences for interpretation.

The force theory predicts that there should be three main causal concepts, CAUSE, HELP, and PREVENT, each associated with a particular configuration of forces[1]. Scenes depicting various configurations of forces are shown in Figure 1; specifically, each scene shows a situation involving a pool of water, a boat with an outboard engine, a bank of fans, and a buoy. The configurations of forces can be made explicit using free-body diagrams, as shown below each scene. As is customary in the construction of free-body diagrams in physics, the forces are shown acting on only one object, the patient. They do not show the location of the affector, only the direction and magnitude of the affector's force on the patient (i.e., **A**). In each of the configurations shown in Figure 1, the patient is associated with a force (i.e., **P**). The force associated with the patient, **P**, can be generated in a number of ways, including from gravity or mechanisms internal to the patient, or it could emerge from the patient's resistance to change because of frictional forces. Copley & Harley (submitted) provide a compelling argument for the view that the force associated with the patient is often best understood as emerging from the patient's position in a 'normal field.' In their account, the normal field gives rise to an object's tendency to fall due to gravity, as well as more abstract tendencies such as entity's tendency to 'grow', 'redden' or straighten'. In Figure 1, the patient's force corresponds to the force generated by the boat's motor. When the patient has a tendency for the end-state, the patient vector, **P**, points in the same direction as the end-state vector, **E**; otherwise, **P** points in a different direction. When the patient and the affector are in concordance, their respective vectors point in the same direction. Finally, the patient entity will approach the end-state when the resultant (sum) of the **A** and **P** vectors, **R**, is in the same direction as the end-state vector, **E.** Importantly, the end-state

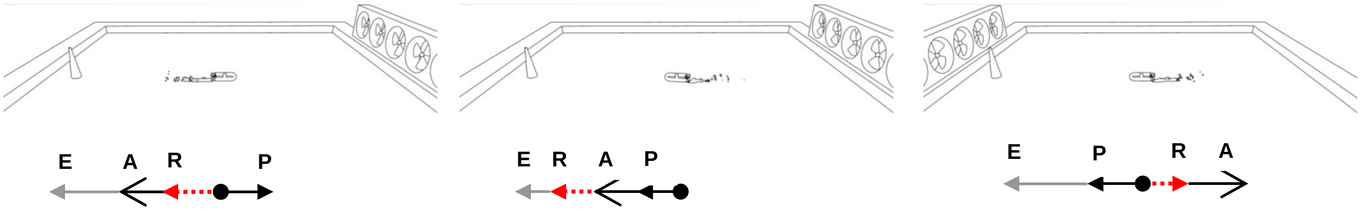
---

[1] ALLOW and ENABLE are associated with the same configuration as HELP, except that the configuration must be derived from a series of PREVENT configurations, as described below.

vector, **E**, is a position vector, not a direction vector. Hence, the length of the end-state vector specifies how close the patient is to reaching the end-state. In a configuration in which the patient is not at the end-state and there are forces acting on the patient that are pushing towards the end-state, the patient will approach the end-state, but a particular configuration of forces and a position vector cannot specify that the patient will actually arrive at the end-state because force and position vectors are fundamentally atemporal. As explained in Copley & Harley (submitted; this volume), that fact that forces can exists without the result necessarily occurring makes forces well suited for explaining ‘defeasible causation’, that is how a causal relation can hold even when the final result does not occur.

A CAUSE configuration is defined as a configuration in which the patient does not have a tendency for the end-state, the affector opposes this tendency, and the resultant is towards the end-state. A HELP configuration is defined as a configuration in which the patient has a tendency for the end-state, the affector is concordant with this tendency, and the result is towards the end-state. A PREVENT configuration is defined as a configuration in which the patient has a tendency for the end-state, but the affector opposes this tendency and the resultant does not point towards the end-state.

Figure 5.1



CAUSE

HELP / ENABLE / ALLOW

PREVENT

Figure 1. Configurations of forces associated with CAUSE, HELP/ENABLE/ALLOW, and PREVENT; **A** = the affector force, **P** = the patient force, **R** = the resultant force; **E** = endstate vector, which is a position vector, not a force. The images above each configuration depict a scene showing a boat, a bank of fans, and a cone. In the CAUSE scene, the boat motors away from the cone, but is pushed back to the cone by the fans. In the HELP scene, the boat motors towards the cone, and the bank of fans push it along in the same direction. In the

Support for the force theory's account of CAUSE, ALLOW, and PREVENT was provided in a series of experiments in which participants categorized 3-D animations of realistically rendered objects with trajectories that were wholly determined by the force vectors entered into a physics simulator. As reported in Wolff (2007), people's descriptions of the events closely matched the predictions of the model.

5.5.2 Composing chains of relations

In addition to specifying how people represent individual relations, the force theory specifies how people may derive new causal relations from chains of relations, a process known as relation composition. For example, given the causal relations *cell phones cause inattention* and *inattention causes accidents,* people may derive the overarching causal relation of *cell phones cause accidents*.

In the force theory, a new overarching causal relation is based on the summative configuration of forces that is derived from the forces in the underlying causal chains. Specifically, the summative configuration of forces uses the same affector force as in the first configuration of forces in the chain and a patient force based on the sum of all of the remaining forces in the chain. While the manner in which the summative configuration of forces is derived is the same across all types of causal chains, the manner in which the causal chain is constructed in the first place depends on whether the chain involves transfers or removals of forces. In transfers, the resultant of the initial configuration of forces in the transfer becomes the affector force in the subsequent configuration of forces. The idea is illustrated in the multiple-collision event shown in Figure 2. In this sequence of events, A begins moving first, it hits B, and B then hits C, sending C over the line. The arrows in Figure 2 indicate the directions of the cars' motion. Cars without arrows are not moving.

Figure 5.2

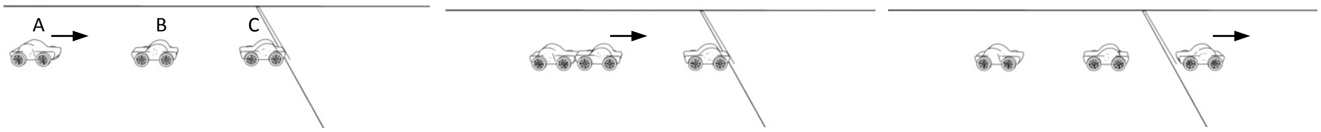


Figure 2. The animation instantiates a CAUSE/CAUSE causal chain in which A causes B and B causes C. A begins moving first. It hits B, sending B into C, which then moves over the line. The animation can be summarized by the sentence *A caused C to cross the line*.

Figure 3 illustrates the forces involved in the animation depicted in Figure 2. On the left side of Figure 3 is a picture of the first frame of the animation. Above B and C are CAUSE configurations of forces. In the first CAUSE configuration, the affector force comes from A and the patient force comes from B, namely its resistance to moving forward due to friction. In the second CAUSE configuration, the affector force comes from B, which is based on the sum of the A and B forces acting on B. In effect, in the second CAUSE configuration, the affector force is the resultant of the A and B forces in the first CAUSE configuration. The patient force in the second CAUSE configuration is C, specifically, its resistance to moving forward. The resultant of the affector and patient forces acting on the third car sends it over the line.

Figure 5.3

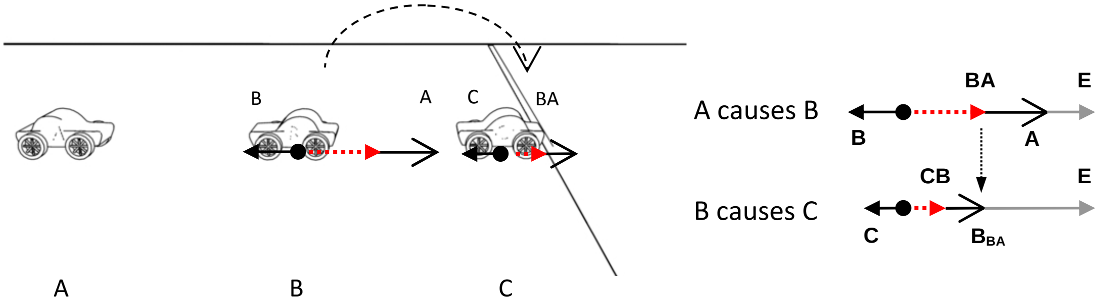


Figure 3. Above B and C are CAUSE configurations of forces. The smaller vectors pointing to the left are the patient vectors acting on B and C (i.e., friction). The longer vectors pointing to the right are affector vectors and the dashed vectors are the resultant vectors. In this sequence of collisions, the resultant vector associated with B becomes the affector vector acting on C. On the right side are two free-body diagrams depicting the same configurations of forces shown on the left, but this time, they are arranged vertically rather than horizontally. In the free-body diagrams, the line pointing down shows how the resultant of the first CAUSE configuration becomes the affector in the second CAUSE configuration.

On the right side of Figure 3 is a pair of free-body diagrams depicting the configurations of forces instantiated in the single frame of the animation. Instead of being arranged horizontally as on the left, the two configurations of forces are arranged vertically. The vertical arrow connecting the resultant vector in the first configuration with the affector vector in the second configuration highlights the fact that the resultant from the first configuration is transferred to the second configuration.

Whereas some causal chains involve transfers of force, other causal chains involve the removal of forces. When a chain involves the removal of a force, the manner in which the causal chain is set up differs. To get a sense for how a causal chain can result from the removal of a force, consider, for example, a chain of PREVENT relations such as A PREVENT B and B PREVENT C. A chain of PREVENT relations is known as a double prevention. At first glance, it may be unclear how a double prevention could be realized because if A first prevents B, B cannot then prevent C because B has already been prevented. The solution to the realization of a double prevention rests in the order in which the prevent relations are realized. In particular, a double prevention can occur if the second prevent relation is realized before the first prevent relation is realized. So, for example, if A prevents B and B prevents C, such a chain can occur if B first prevents C before A prevents B. The intuition behind this can be illustrated with a real world example. Consider pulling a plug to allow water to flow down a drain. Such a sequence of PREVENTs begin with the plug (B) preventing the water (C) from draining (that is the second relation in the double prevention). Then, someone (A) prevents B by pulling the plug, that is, by removing B’s force on C. Note that when A pulls B, A opposes not just the force associated with B, but also the force associated with C, that is, the resultant of the **B** and **C** forces (the plug and the water). Thus, in the case of double prevention, the resultant of the second premise (**CB**), which is computed first, serves as the patient vector in the first premise ($\mathbf{B}_{\mathbf{CB}}$).

The way forces are transmitted in a double prevention can be illustrated in a different way based on the chain depicted in Figure 4.

Figure 5.4

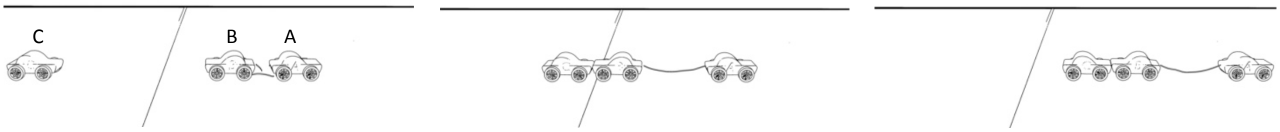


Figure 4. The still frames depict key stages in a PREVENT / PREVENT chain. First, C attempts to cross the line but is prevented by B. Then, A pulls B away from C with a rope, preventing B from preventing C. With the removal of B, C crosses the line.

In the beginning of the animation depicted in Figure 4, C approaches the line. B then approaches C and prevents it from crossing the line. The middle panel shows A pulling B away. In the panel on the far right, with the removal of B, C crosses the line. The forces involved in the animation shown in Figure 4 are depicted in Figure 5.

Figure 5.5

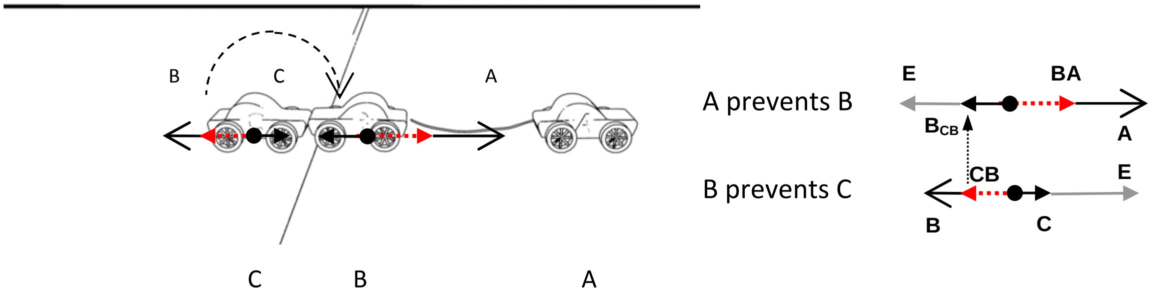


Figure 5. The scene depicts the configuration of forces instantiated in a PREVENT / PREVENT causal chain. The smaller vectors pointing left and right are patient vectors, while the longer vectors are affector vectors. The dashed vectors are resultant vectors. On the right side are two free-body diagrams depicting the same forces shown on the left, arranged vertically. In the diagrams, the vector **E** is the position vector pointing to the end-state, which, in the animation on the left, is the area on the right side of the line. Note that in a double prevention, the resultant vector of **B** and **C** becomes the patient vector in the interaction between **B** and **A**.

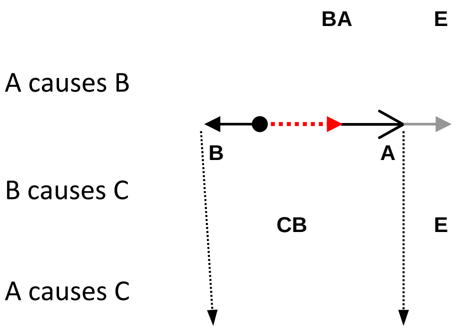
Figure 5 illustrates the forces involved in the animation depicted in Figure 4. On the left side of Figure 5 is a frame from one of the middle frames of the animation. Above B and C are PREVENT configurations of forces. In the first PREVENT configuration, the affector force comes from B and the patient force comes from C. In the second PREVENT configuration, the affector force comes from A while the patient force comes from the sum of B and C. Hence, the patient in the first PREVENT configuration is the resultant of the B and C forces in the second PREVENT configuration, which occurs first.

On the right side of Figure 5 is a pair of free-body diagrams depicting the same configuration of forces shown in the frame of the animation, this time arranged vertically. The diagram above depicts the configuration of forces acting on B, while the diagram below depicts the configuration of forces acting on C. The vertical arrow connecting the resultant vector in the second configuration with the patient vector in the first configuration highlights the fact that the resultant is transferred from one configuration to the other. As discussed above, in chains of PREVENT relations, the resultant of the second PREVENT configuration serves as the patient vector in the first PREVENT configuration.

5.5.3 Composing causal relations

As discussed earlier, the process of composing causal relations involves constructing an overall summative configuration of forces based on all of the configurations in the chain. Whether the chain involves the transfer or removal of a force, the manner in which a summary configuration is derived is the same. Figure 6 shows how a summary configuration is derived from a chain of two CAUSE relations. As depicted in Figure 6, the affector in the summary configuration is the affector from the first relation (**A**); the end-state is based on the end-state vector in the last relation (**E**); and the patient in the summary configuration is derived by summing all of the remaining forces in the causal chain (**B**+**C**).

Figure 5.6



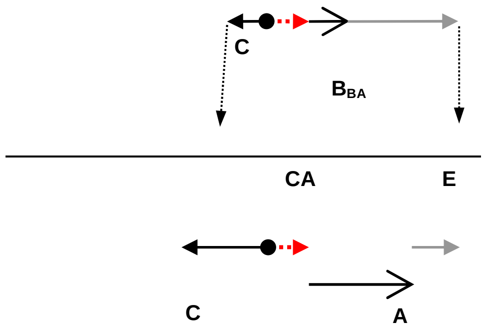


Figure 6. The affector force in the summary configuration, **A**, is the affector force in the first relation, **A**. The end-state in the summary configuration is the end-state vector from the last relation. The patient force in the summary configuration, **C**, is based on the vector addition of all of the patient forces in the chain (**B** and **C)**.

5.6 Explaining various causal phenomena in terms of the force theory

The force theory offers a comprehensive account of various causal phenomena that previously could not be explained by any single theory of causation.

5.6.1 ALLOW relations

The force theory offers an account of how people represent ALLOW relations. Following McGraph (2003), we propose that ALLOW relations are necessarily based on double preventions (see also Barbey & Wolff 2007, Wolff et al., 2010). In the simplest case, ALLOW relations involve removing a force that was originally preventing it from happening. An example is given in the animation depicted in Figure 4: when A pulls B away from C, it allows C to cross the line. This account of ALLOW was supported in a set of studies described in Wolff et al. (2010). Participants saw animations like the one shown in Figure 4 and were asked to choose one of several possible descriptions. Across a range of double preventions, participants indicated that the relationship between the first and last entities in a double prevention instantiated an ALLOW relation[2]. A more complex kind of ALLOW relation involves what Wolff et al. (2010) refer to as a *virtual force*. In this kind of allowing, the initial prevent relation is threatened rather than actualized. ALLOW relations involving virtual forces are easy to imagine. Consider, for example, a car that is about to cross a line and a second car moving towards the first car that will prevent the first car from crossing the line if the cars collide. At the last moment, the second car stops and the first car is 'allowed' to cross the line. In such a scenario, the second car threatens to prevent the first car from crossing the line, but this threat is not actualized, so the force that would have prevented the first car from crossing the line is a virtual force. Evidence showing that people interpret such events as instantiating ALLOW relations was found in a set of studies by Wolff et al. (2010). Participants saw events like the one described above. As predicted, participants endorsed the descriptions that said that the second car allowed the first car to cross the line.

5.6.2 Causation by omission

The force theory offers an account of how people understand causation by omission. As noted earlier, causation by omission is causation in which an absence of an influence brings about an effect, as in *The absence of nicotine causes withdrawal* or *Lack of light causes depression*. The phenomenon of causation by omission is especially problematic for production theories of causation that define causation in terms of the transmission of energy or force for the obvious reason that there cannot be transmission from an absence. In the force theory, however, causation

---

[2] We further propose that the concept of ENABLE is closely related to that of ALLOW. They both entail double prevention but differ slightly in the manner in which the particular PREVENT relations are instantiated.

can be based not only on transmission but also on removal of an actual force or threat of a virtual preventive force (Wolff et al., 2010). Such effects often occur in many instances of double prevention. Consider again the example of double prevention involving the Dutch boy who pulled his finger from the dam and consequently flooded the town. As noted earlier, we could say that the boy caused the flooding by preventing his prevention of the leak. We could also describe the same scenario in terms of an absence, specifically, that *The absence of the Dutch boy's finger caused the flooding of the town*. Another example of double prevention discussed earlier was that of pulling a plug, causing or allowing water to flow down a drain. This scenario could also be described as resulting from an absence: *The absence of the plug caused/allowed the water to flow down the drain*. Thus, the proposal is that causation by omission is always embedded within a double prevention, specifically, that causation by omission involves the relationship between the second and third entities in a chain of prevent relations. In double preventions, the second entity is removed by the first entity, so the relationship between the second and third entities concerns what happens to the third entity in the absence of the second. The double prevention shown in Figures 4 and 5 provides yet one more example of causation by omission. As noted earlier, Wolff et al. (2010) found that people viewed the relationship between the first and last cars in this causal chain as instantiating an ALLOW relation, as evidenced by their willingness to endorse the statement *A allowed C to cross the line*. Wolff et al. (2010) also found that people were willing to view this causal chain as an instance of causation by omission, as evidenced by their willingness to endorse the statement *The absence of B's influence allowed C to cross the line*. Despite arguments to the contrary (Schaffer 2000, Schulz et al. 2007, Woodward 2006), a force-based approach is capable of representing causation by omission.

The concept of causation by omission may underlie the notions of triggering and releasing. Consider, for example, a situation analyzed by Kistler (2006a) in which a metal catch is released, allowing/causing a spring to move a weight. As noted by Kistler (2006a), release of the metal catch does not contribute any energy to the movement of the weight; the movement of the weight is produced only by the energy stored in the spring. Rather, the release of the metal catch involves the removal of a force that allows another force to bring about an effect. The situation can be described without having to make explicit reference to an absence: we can say ‘release of the metal catch caused the spring to move the weight.’ The example illustrates how causation by omission might be embedded in ordinary statements of causation, without explicit reference to the omission (see also Schaffer, 2000). Moreover, given that the example above seems to exemplify a type of ‘triggering’, it may be that triggering events are causal events in which there is implicit causation by omission.

5.6.3 Synonymy relations

As discussed in Wolff et al. (2010), dependency theories make predictions about how different kinds of causal expressions are related to each other. For example, according to Cheng & Novick’s (1992) probabilistic contrast model, a CAUSE relationship implies that the probability of an effect in the presence of a case, $P(E|C)$, is greater than the probability of the effect in its absence, $P(E|\neg C)$, that is, $P(E|C) > P(E|\neg C)$. A PREVENT relationship implies the opposite, $P(E|C) < P(E|\neg C)$. In Cheng and Novick’s theory, the statement *¬C causes E*—causation by omission—implies that the probability of E given ¬C, $P(E|\neg C)$, is greater than the probability of E given C, $P(E|C)$, that is, $P(E|\neg C) > P(E|C)$, which is the same inequality as that for PREVENT. In other words, the inequality associated with *¬A causes B* is the same as that for *A prevents B*. The probabilistic contrast model also predicts that *C causes ¬E* is the same as *A prevents B*, since if *C causes ¬E*, then $P(\neg E|C) > P(\neg E|\neg C)$, which entails $P(E|C) < P(E|\neg C)$. A more recent dependency model, the causal model theory, makes the same predictions about the relationship between negated CAUSE claims and PREVENT relations (Sloman, Barbey & Hotaling, 2009). A key feature of the causal model theory, which is a Bayesian network theory of causation, is that it defines causal relations in terms of different causal structures, which allows this theory to represent not only CAUSE and PREVENT relations but also ALLOW relations. As described in Wolff et al. (2010), the causal model theory predicts that negated ALLOW relations are related to PREVENT. Specifically, the model predicts that ¬*A allows B* and *A allows* ¬*B* both entail PREVENT. Importantly, in both the probabilistic contrast model and the causal model theory, there are no asymmetries in the degree to which negated CAUSE relations or ALLOW relations (in the case of the causal model theory) are related to PREVENT relations: ¬*C causes E, C causes* ¬*E*, ¬*A allows B*, and *A allows* ¬*B* all imply or entail *A prevents B* to an equal degree.

The force theory, in contrast, predicts that PREVENT relations are more similar to certain kinds of CAUSE and ALLOW expressions than others depending on whether the cause or effect is negated. As discussed in Wolff et al. (2010), according to the force theory, the relationship between negative and positive causal statements can be explained in terms of sub-chains within an overall causal chain. For example, the statement *Lack of B allows C* implies a chain of two PREVENT relations, i.e., A PREVENTS B and B PREVENTS C. If we want to know what would happen in the *presence* of B, we eliminate the first PREVENT relation and evaluate the chain with respect to the remaining PREVENT relation. In this case, the remaining PREVENT relation would be *B prevents C*; hence, the theory predicts that the claim *Lack of B allows C* implies (*Presence of) B prevents C*. The strategy for deriving this implication can be generalized to other types of omission. For example, *Lack of A causes D* implies a chain of two PREVENT relations (A PREVENT B, B PREVENT C) followed by one CAUSE relation (C CAUSE D). To determine the relation implied when A is present rather than absent, we remove the initial PREVENT relation from the chain and compose the remaining two relations, B PREVENT C and C CAUSE D. As discussed earlier, the process of relation composition involves adding the patient vectors in the component relations. As noted in Wolff et al. (2010), when the magnitudes of the patient vectors are exhaustively varied[3], the composition of PREVENT and CAUSE leads to a PREVENT relation 37% of the time and an undefined summary configuration 63% of the time. In the case of *Lack of A allows D*, the resulting sub-chain is a PREVENT relation 100% of the time. Hence, according to the force theory, the expression *Lack of A causes D*, does, in fact, imply *B prevents D* but to a weaker degree than *Lack of A allows D*. As described in Wolff et al. (2010), the same process can be used to determine the degree to which *A causes lack of C* and *A allows lack of C* imply *A prevents C*. The result is that the force theory predicts A PREVENT B is more strongly associated with A CAUSE ¬B than with ¬A CAUSE B. In addition, the force theory predicts that A PREVENT B is more strongly associated with ¬A ALLOW B than with A ALLOW ¬B.

In Wolff et al. (2010), these predictions were tested by having participants read causal claims and then having them choose the most synonymous paraphrase. The paraphrases differed with respect to the verb (*cause, allow, prevent*) and whether the cause or effect term was expressed as present or absent. The results were as predicted by the force theory. When people were presented with *A prevents B*, they chose *A causes ¬B* (36%) and *¬A allows B* (34%) more often than they chose *¬A causes B* (15%) and *A allows ¬B* (5%). In sum, the force theory correctly predicted certain

---

[3] All possible combinations of patient vectors can be calculated by using integral calculus or via a iterative procedure for exploring the space of possible magnitudes,

correspondence relationships, including asymmetrical ones, among causal statements. Other theories of causation can neither predict nor account for such correspondences.

5.7 Challenge: Is the force theory limited to only physical relationships?

In the previous section, we demonstrated how the force theory accounts for phenomena that dependency theories cannot, namely causation by omission and double prevention (Schaffer 2000, Schulz et al. 2007, Woodward 2006). In addition, the force theory is uniquely able to account for asymmetries in the way various causal expressions are related to each other. Despite these successes, a causal pluralist might be willing to grant force-like representations in the case of simple physical causation but hold off on the use of such representations in the case of abstract causation. Any reluctance about the use of forces in the representation of abstract causation would be understandable given the obvious fact that abstract causal relations lack actual physical forces. For example, what are the physical forces in causal relationships such as *Tax cuts cause economic growth* or *Competition prevents inflation*? Moreover, even in cases where the underlying mechanisms are physical, it has been empirically established that people typically have no idea about how the mechanisms are grounded in actual forces in the world (Rozenblit & Keil, 2002). For example, when told that *Animal products cause cholesterol* or that *Diabetes causes foot problems*, people most likely have no idea how these causal relations might ultimately physically instantiated. So even in cases of certain kinds of physical causation, the underlying representations may not be based on forces. Given these observations, it may be that people’s understanding of relatively simple physical causation is based on forces, but that people’s understanding of abstract causation, or complex types of physical causation, are based on some other type of representation, such as statistical or counter-factual dependency.

According to Talmy (1988), however, force dynamics is not limited to physical causation. Talmy (1988) held, for example, that intentions and desires could be treated as analogous to physical forces. Empirical support for this idea was provided by Wolff (2007) with materials in which physical forces were replaced with intentions or desires. It was found that force dynamics provided an excellent account of how people described causal situations even when all of the ‘forces’ were abstract. But perhaps the scenes examined in Wolff (2007) were still relatively concrete. The scenes all showed two people interacting with each other, which is far simpler than the type of causation implied by the statement *Lack of education causes poverty*. Perhaps for these very abstract cases of causation, people may rely on a dependency concept of causation.

Synonymy relationships between different causal expressions can be used to test this possibility. As discussed earlier, Wolff et al. (2010) found that the expression *A prevents B* was more similar in meaning to *A causes not B* and *not A allows B* than it was to *not A causes B* or *A allows not B.* These correspondence asymmetries are not predicted by dependency or other process theories. Perhaps, then, we might find that the magnitude of this asymmetry varies with the type of causation under consideration. Specifically, we might find that these asymmetries exist when people make synonymy judgments about physical causation, but not about abstract causation. The asymmetries would be expected to disappear if people thought about abstract causation in terms of dependency. This possibility was tested in the following experiment.

5.7.1 Experiment: Synonymy relationships in physical and abstract causation

Participants (N = 70) were shown pairs of causal sentences and asked to rate on a 0-to-10 scale the degree to which two sentences were related to each other, with 0 indicating that the two sentences were completely unrelated and 10 indicating that the two sentences were highly related.

Participants were instructed to give high ratings when they felt that the two sentences had the same meaning or if one sentence entailed the other. The sentence pairs were constructed from a list of 84 causal statements obtained from Google searches. One third of the sentences were CAUSE sentences, another third were ALLOW sentences, and the remaining third were PREVENT sentences. In each third, half of the sentences described physical relationships and the remaining sentences described abstract relationships[4]. Examples of physical prevention included *Heating prevents condensation*, *Salt prevents ice*, and *Dust prevents adhesion*. Examples of abstract prevention included *Stress prevents clear thinking*, *Tourism prevents unemployment*, and *Accountability prevents corruption*. Each test sentence was paired with all possible combinations of the sentence that could be formed by replacing the causal verb with another causal verb and with either the cause or effect nouns negated. The resulting 336 possible pairs of sentences were divided across 14 conditions with each participant rating just 24 of the possible pairings.

Because process and dependency theories make different predictions about how PREVENT sentences are related to other causal expressions, the follow analyses focus on participants' ratings of PREVENT sentences, specifically how *A prevents B* was related to *A causes lack of B* (C¬), ¬*A causes B* (¬C), *A allows* ¬*B* (A¬), and ¬*A allows B* (¬A). On the left side of Figure 7 are the average relatedness ratings for physical relationships and on the right side are the average relatedness ratings for abstract relationships.

Figure 5.7

---

[4] The relative abstractness of the causal relations was determined by a separate group of participants (N = 48), half of whom rated the relations on a 0-to-10 scale with respect to their abstractness and the other half of whom rated the relations on a 0-to-10 scale with respect to physicality.

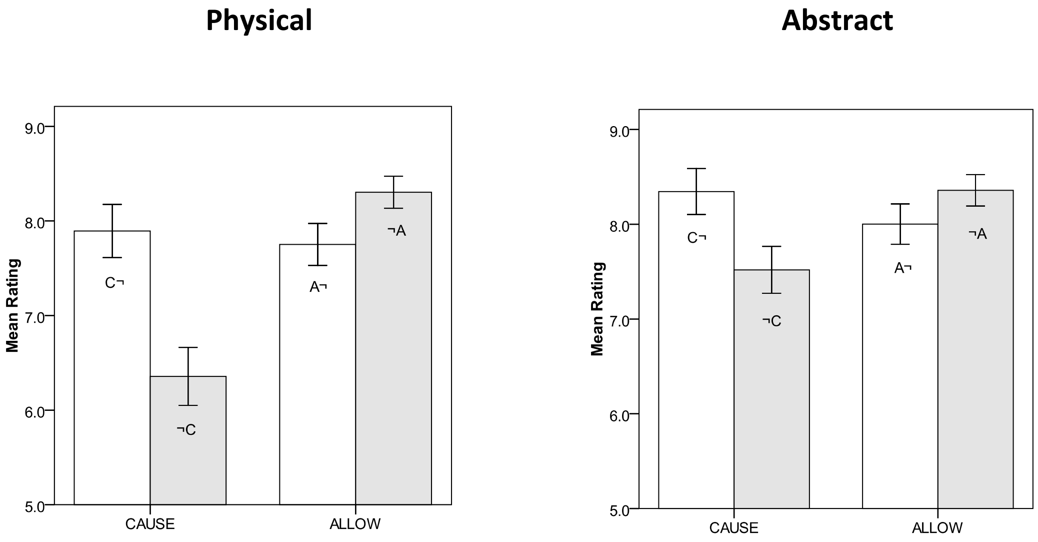


Figure 7. People’s ratings of the relatedness of PREVENT sentences (A prevents B) with CAUSE¬ (C¬),¬CAUSE (¬C), ALLOW¬ (A¬), and ¬ALLOW (¬A) sentences for physical and abstract relationships, along with standard errors of the mean.

The relatedness ratings in this experiment effectively replicated the multiple choice decisions in Wolff et al. (2010). In particular, ignoring the distinction between physical and abstract, PREVENT sentences were rated as more related to CAUSE¬ sentences than ¬CAUSE sentences, t(69) = 6.16, p < .001; and PREVENT sentences were also rated as more related to ¬ALLOW sentences than ALLOW¬ sentences, t(69) = 3.25, p < .01. Importantly, these contrasts were still significant when the data was analyzed according to physicality of the relation. PREVENT sentences were rated as more related to CAUSE¬ sentences than ¬CAUSE sentences for both physical, t(69) = 5.85, p < .001, and abstract relations, t(69) = 3.13, p < .01; and PREVENT sentences were rated as more related to ¬ALLOW sentences than to ALLOW¬ sentences for both physical, t(69) = 3.22, p < .01, and abstract relations, t(69) = 2.21, p < .05. Thus, in terms of individual contrasts, similar relatedness asymmetries were found for both physical and abstract causal relations. This is in contrast to the hypothesis that abstract causal relations might be understood in terms of dependency and physical causal relations in terms of forces, in which case these asymmetries should not have been present for abstract causal relations.

5.8 Causation as a prototype concept

Up to this point, I have focused on what is arguably the simplest as well as most common type of causal pluralism: the idea that the notion of causation might be associated with two types of representations, production causation and dependency causation. However, another school of causal pluralism does not limit the notion of CAUSE to just two concepts. According to Cartwright (2004), causation comprises a wide range of relations that are reflected in words like *feed, open, poison, enrich, clog, attract* and *compress*. Cartwright (2004) refers to these specific kinds of causation as ‘thick’ causal concepts; in contrast, the causation encoded in the verb *cause* exemplifies a ‘thin’ causal concept. She further argues that thin causal concepts can be very useful, but they are ultimately derived from thick causal concepts. Thus, for Cartwright (2004), the concept of CAUSE can be many things depending on which set of thick concepts is relevant in the context. Her position is similar to that of Psillos (2008), who likens causation to the common cold: a loose collection of symptoms that may be produced in any number of ways. Likewise, Lakoff & Johnson (1999) view causation as a prototype concept at the center of a category in which there are many members, but no commonality across all of the members. According to prototype theory, each concept of cause has its own logic (Lakoff & Johnson, 1999). However, as shown above, the results of our asymmetry experiment provide no evidence for Lakoff & Johnson’s (1999) conjecture. The logic governing synonymy relationships between different expressions of causation appears to be essentially the same for both physical and abstract causal relations. Moreover, there has been no demonstration in the literature of transitive reasoning depending on the type of causal relations being composed. To date, then, the logic underlying causal relations appears to be the same for the many types of relations that are viewed as causal.

5.9 Conclusions

The pluralist position is motivated by the failure of process and dependency theories to account for the various phenomena associated with causation. In this chapter I have argued that the move towards multiple types of causation is not necessary once we consider production models that go beyond simple transmission of energy, namely the force theory. The force theory is able to handle phenomena such as causation by omission and double prevention, which have been highly problematic for production theories as well as phenomena such as causal redundancy and transitivity, which are problematic for dependency theories. In addition, the force theory is able to address phenomena that have not been adequately addressed by either of the two main classes of theories, such as the distinction between CAUSE and ENABLE, causation by omission, and synonymy relationships between different types of causal relations. The success of the force theory suggests that a common notion of causation might ultimately be possible.