Wolff, P., Hausknecht, M., & Holmes, K. J. (2011). Absent causes, present effects: How omissions cause events. In J. Bohnemeyer & E. Pederson (Eds.), *Event representation in language and cognition* (pp. 228-252). Cambridge, UK: Cambridge University Press.

11 Absent causes, present effects

How omissions cause events

Phillip Wolff, Matthew Hausknecht, and Kevin Holmes

1 Introduction

Causal relationships range from the physical to the abstract: from friction causing heat to stress causing forgetfulness. This broad spectrum of relationships motivates the question of what all causal relationships have in common. One approach has been to specify the conditions for causation in terms of the occurrence or non-occurrence of events or states, with no regard to processes that produce these events or states. Because these theories specify causation in terms of the effects of causation, they will be referred to as *outcome* theories. Outcome theories typically describe the conditions for causation in terms of probabilities, counterfactuals, first-order logic or mental models. An alternative approach specifies the conditions for causation in terms of the processes that bring about outcomes; such accounts will be referred to as *process* theories. Process theories typically specify the conditions for causation in terms of the transmission of energy and force or their analogs in the social and psychological domains, for example, intentions and social pressure.

The two kinds of theories sometimes address different questions about causation, making them, in some sense, complementary. However, they contrast sharply on the question of what counts as a causal event, in particular, the phenomenon of causation by omission. Causation by omission occurs when the absence of an influence brings about an effect. We say, for example, *Not watering the plant caused it to wilt* or *Lack of clean air causes dizziness.* Outcome theories view causation by omission as a fully legitimate kind of causation, a position that we will support in this chapter. Outcome theories hold this position because the criteria for causation in these theories do not depend on the underlying processes that give rise to events; they do not need to explain how the absence of an influence could cause something to happen. Process theories, on the other hand, deny that causation by omission is causation in the fullest sense of the concept. They are led to this position because, as currently formulated, process theories define causation in terms of the transmission of force, and plainly nothing can be transmitted by an absence.

In this chapter we will argue for the legitimacy of causation by omission, but against the view that causation is defined in terms of outcomes. We will argue instead that causation is fundamentally understood as specified in process theories. Although process theories, to date, have denied the legitimacy of causation by omission, we will argue that the rejection of this kind of causation does not necessarily follow from this approach to causation. We will then show how the phenomenon of causation by omission can be handled in process theories.

We begin with a brief discussion of outcome theories and how they specify causation by omission. We also note several key challenges for outcome theories. We then provide a brief discussion of various process theories, including the two key challenges that have been raised against these theories. One challenge was originally raised by Hume (1978 [1739]), namely that accounts of causation based on notions such as force are circular. We address Hume's criticism with both logical and empirical evidence. We then take up the remaining challenge of explaining causation by omission in terms of a process. In particular, we will present a new process theory of causation, the **force theory**, which explains how causation by omission and commission can be given a unified characterization. In the last section, we describe empirical evidence in support of this new account of causation.

2 Defining causation in terms of outcomes

According to outcome theories, causal relations are specified in terms of the occurrence or non-occurrence of events or states, without regard to the nature of the process that produced those events or states (Ahn and Kalish 2000). Because the mechanism is left unspecified, such theories, in effect, avoid the problem raised by causation by omission since they do not attempt to specify the manner in which an effect might be caused by an absence. There are several classes of theories that specify causation in terms of outcomes. Here we quickly review some of these theories and how they account for causation by omission.

2.1 Probability raising models

According to probability raising models, a cause can be defined as an event that changes the probability of another event. The concept of CAUSE, in particular, as opposed to the concept of PREVENT, is usually associated with probability-changing events in which the probability of an effect in the presence of a cause, $P(E|C)$, is noticeably greater than the probability of an effect in the absence of a cause, $P(E|\neg C)$ (Cheng and Novick 1991, 1992). When the probability of the effect in the presence of a cause is greater than in its absence, it is often said that the cause "raises the probability of the effect." For example, the probability of traffic jams is greater in the presence of construction than in its absence, so according to the probability criterion of causation, construction causes traffic jams. In these theories, omissions are treated like commissions. For example, in probability-raising models, the statement *Lack of water caused the plant to wilt* – causation by omission – would entail that the probability of E given $\neg$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). Thus, causation by omission does not raise problems for these kinds of models.

2.2 *Counterfactual theories of causation*

Counterfactual theories offer a second type of outcome theory (Lewis 1973, 2000; see also Spellman and Mandel 1999). According to counterfactual theories, 'C causes E' holds if it is the case that if C had not occurred, E would not have occurred. Because this criterion is stated in terms of the occurrence and non-occurrence of events, it can be adapted to causation by omission. As argued by McGrath (2005), the meaning of the causal claim 'Not C causes E' would presumably map onto the conditional "if C had occurred, E would not have occurred" (e.g., *If watering had occurred, wilting would not have occurred*). As with probability raising models, counterfactual theories handle causation by omission without difficulty.

2.3 *Mental model theory of causation*

A third type of outcome theory is Goldvarg and Johnson-Laird's (2001) mental model theory. The model theory goes beyond other theories in characterizing not only CAUSE and PREVENT, but also distinguishing these two notions from ALLOW. According to the mental model theory, the notions of CAUSE, ALLOW, and PREVENT are associated with different combinations of possible co-occurrences. For example, a CAUSE relation is associated with a set of co-occurrences in which A is present and B is present (a b), A is absent and B is present ($\neg$a b), and A is absent and B is absent ($\neg$a $\neg$b). Applying NOT to the antecedent or consequent flips the states of affairs of the antecedents and consequents (respectively) in all of the possible co-occurrences. Thus, the meaning of "Not C causes E" would be given by a set of co-occurrences ($\neg$a b), (a b), and (a $\neg$b). As with the previous outcome theories, causation by omission is handled quite easily by the mental model as simply a different set of possible co-occurring states.

2.4 *Causal Bayesian network theories*

A fourth type of outcome theory is represented by causal Bayesian network theories of causation. In causal Bayesian networks, variables are connected

to one another with 'arcs', as in A $\rightarrow$ B. Each arrow in a causal Bayesian network is associated with a set of conditional probabilities in conjunction with assumptions about the effect of actual or hypothetical interventions (Woodward 2003, 2007; Sloman 2005; Schulz, Kushnir, and Gopnik 2007). A recent proposal by Sloman *et al.* (2009), the causal model theory, shows how a Bayesian network approach to causation might be applied to the representation of CAUSE, ALLOW, and PREVENT, including causation by omission. Sloman *et al.* (2009) frame their theory in terms of structural equations, which represent a particular way of instantiating a graph with arrows. For example, the graph A $\rightarrow$ B can be instantiated in a structural equation such as B := A (Sloman *et al.* 2009; see also Hitchcock 2001). According to their theory, the concept of ALLOW is associated with a different structural equation, namely, the claim *A allows B* is represented as B := A and X, in which the variable X is an assessor variable. The claim *A prevents B* is represented by the equation B := $\neg$A, which also represents causation by omission, that is, claims such as *Not-A causes B*. As with the other outcome theories, causation by omission falls out naturally from causal Bayesian network theories such as Sloman *et al.*'s (2009) causal model theory.

2.5 *Challenges for outcome theories*

Special assumptions are not required for the representation of causation by omission in outcome theories; hence, these theories provide a unified account of different kinds of causation. In addition, these theories define causation in terms that are relatively unambiguous rather than relying on constructs that are as undefined as the concept they seek to explain. Despite these strengths, there are several properties of causation that outcome theories cannot account for: temporal priority, mechanism, and spatio-temporal contiguity. Temporal priority is the property that causes must precede or occur simultaneously with their effects. In outcome theories, temporal priority needs to be stipulated (Hitchcock 2002) because the representations used in outcome theories do not require that the causal factors occur in a certain chronological order. In probability-raising theories, for example, a correlation between variables A and B can exist regardless of their temporal order. Outcome theories also do not explain the importance of mechanism in people's causal knowledge, that is, why people expect non-contiguous events to be connected via a chain of intervening events. Correlations and counterfactual dependencies, for example, do not depend on the presence or absence of intervening links. Finally, outcome theories cannot motivate why spatio-temporal contiguity should be relevant to causation. Some researchers have suggested that spatio-temporal contiguity may be important to causation because of an innate bias to attend to these features (Schulz *et al.* 2007; see also Gopnik *et al.* 2004; Leslie and Keeble 1987), but such a proposal really offers no explanation, or actually concedes the point: if attention to spatial and temporal contiguity is innate, it must be because it is an important property of causation. Outcome theories do not explain why this should be the case.

In contrast to outcome theories, process theories are able to account for the properties discussed above. First, the property of temporal priority is motivated in process theories because a result can only occur after there has been an exchange of energy or force. Second, process theories motivate the need for mechanism. Forces impose a relatively "local" level of granularity. This local level of granularity implies that for indirect causal relationships, that is, for causal relationships between non-contiguous events, there must be a sequence of intermediate links, each contiguous to the next (Russell 1948). Thus, process theories entail the existence of causal mechanisms. This assumption has been strongly supported by work in psychology (Ahn and Bailenson 1996; Ahn and Kalish 2000; Ahn *et al.* 1995; see also Bullock, Gelman, and Baillargeon 1982; Shultz 1982).

A third property of causation that is readily explained by process theories but not by outcome theories is the importance of spatio-temporal contiguity. The importance of spatial and temporal contiguity has been repeatedly observed in standard "billiard ball" type events in which one object hits and launches another into motion (Michotte 1963 [1946]). Much research has shown that the impression of causation is greatly diminished if the first object in the sequences does not make physical contact with the second object (Lesser 1977; Michotte 1963 [1946]; Oakes 1994; Spelke, Phillips, and Woodward 1995; Thinès, Costall, and Butterworth 1991), or if the second object moves after a significant delay (Kruschke and Fragassi 1996; Leslie 1984; Oakes 1994; Spelke *et al.* 1995; for a review, see Scholl and Tremoulet 2000). Process theories are able to explain the importance of spatio-temporal contiguity because they define causation in terms of forces and physical forces require spatio-temporal contiguity (Wolff 2008).

In sum, outcome theories offer unified accounts of both causation by commission and causation by omission, but they cannot explain the properties of temporal priority, mechanism, and spatio-temporal contiguity. Given that these properties are central to people's concept of causation, process approaches to causation remain viable, despite some limitations. Below, in addition to reviewing various process theories, we will propose how the limitations of process theories might be overcome.

3 Defining causation in terms of process

Process theories begin with the assumption that causation in the mind can be traced to the physical world (Dowe 2007; Wolff 2007). In almost all process theories described to date, causation involves a transmission of these physical

quantities. For example, according to Aronson's (1971) transference theory, causation implies contact between two objects in which a quantity possessed by the cause (e.g., velocity, momentum, kinetic energy, heat, etc.) is transferred to the effect. Another transference theory is proposed by Fair (1979), who holds that causes are the source of physical quantities, energy, and momentum that flow from the cause to the effect. According to Salmon's (1994, 1998) invariant quantity theory, causation involves an intersection of world lines that results in the transmission of an invariant quantity. The proposals of Aronson, Fair, and Salmon come from philosophy. Similar proposals from psychology have been termed 'generative theories' of causation. According to Bullock *et al.* (1982), adults believe that causes bring about their effects by a transfer of causal impetus. Shultz (1982) suggests that causation is understood as a transmission between materials or events that results in an effect. According to Leslie (1984), physical causation is processed by a 'theory of bodies' that schematizes objects as bearers, transmitters, and recipients of a primitive notion of force.

A recent proposal from philosophy breaks from earlier process models in not requiring that the transmission occur in only one direction. According to Dowe's conserved quantity theory (2000), there are two main types of causation: persistence (e.g., inertia causing a spacecraft to move through space) and interactions (e.g., the collision of billiard balls causing each ball to change direction). Causal interactions occur when the trajectories of two objects (essentially, Salmon's 'world lines') intersect and there is an *exchange* of conserved quantities (e.g., an exchange of momentum when two billiard balls collide). Unlike earlier theories, exchanges are not limited to a single direction (i.e., from cause to effect); however, even in the case of an exchange, at least part of the interaction involves a transmission from the cause to the effect.

As discussed earlier, one of the strengths of process theories is that they explain why the properties of temporal priority, mechanism, and spatio-temporal contiguity are important to causation. However, these theories have been subject to at least two key criticisms. One of the criticisms was originally raised by Hume (1978 [1737]) and concerns the value of defining causation in terms of abstract notions such as force. The second criticism concerns the ability of process models to address causation by omission and related phenomena. In the following sections, we will address the challenges raised by these criticisms. With respect to Hume's challenge, we will discuss evidence that force is not necessarily abstract but rather something physical that people are able to apprehend through bodily senses. With respect to the problem of causation by omission, we will describe a new process theory of causation that explains how causation by omission can be specified in terms of causal processes.

4 Addressing Hume's challenge

An important potential criticism of the process theories of causation was originally raised by Hume (1978 [1739]), but is often repeated in current discussion of theories of causation (Cheng and Novick 1991, 1992; Cheng 1997; Schulz, Kushnir, and Gopnik 2007; Woodward 2007). According to Hume, defining causation in terms of force is circular because the notion of force cannot be defined without reference to causation itself. Further, the reason why the notion of force has no independent standing from the notion of cause is because forces cannot be directly observed. Below we provide evidence against Hume's criticism by showing that forces can be directly apprehended by our sensory systems. Second, we discuss behavioral and linguistic evidence that the notion of force does not depend on causation. Specifically, we review evidence that people are able to distinguish between events on the basis of forces. We conclude that people can represent the notion of force independent of the notion of CAUSE, and so Hume's argument that definitions of causation based on force are circular can be rejected.

As argued by White (1999, 2006), causal understanding might begin in the somatosensory system, which includes kinesthesia, the sense of the body's position, weight, or movement, as well as skin pressure sensors. The somatosensory system allows for the direct detection of dynamic properties, such as mass, force, and energy (e.g. heat). In the motor control literature, this idea is reflected in what is referred to as an 'internal model,' that is, a representation that specifies the dynamics of the environment (Reinkensmeyer, Emken, and Crammer 2004; Kurtzer, Herter, and Scott 2005; Davidson and Wolpert 2004; Conditt, Gandolfo, and Mussa-Ivaldi 1997; Hinder and Milner 2003; Imamizu, Uno, and Kawato 1995; Kawato 1999; Ohta and Laboissière 2006; Milner and Franklin 2005; Papaxanthis, Pozzo, and McIntyre 2005; Shadmehr and Mussa-Ivaldi 1994). Our internal models are at work when, for example, we over-lift a suitcase that we think is full when it is, in fact, empty (Reinkensmeyer *et al.* 2004).

One of the classic studies on internal models was conducted by Shadmehr and Mussa-Ivaldi (1994). Participants were instructed to move their hand from one point to another while holding onto a handle that was part of a robotic arm (i.e., a manipulandrum). The robotic arm was programmed to generate forces that pushed the person's hand away from the target location. With repeated practice, people learned how to overcome the pressure of the robotic arm and to reach straight for the intended target. The key finding was the appearance of an after-effect once the force field (robotic arm) was removed: people's arm trajectories were distorted in the opposite direction of the previously applied force. The result can be explained as due to the formation of an internal model of the dynamics of the environment that continued to affect motion even after the dynamics of the environment returned to normal. Similar findings have been observed in conditions of microgravity, that is, when people are asked to reach for targets when they are in weightless conditions of parabolic flight (Papaxanthis *et al.* 2005). Changes in the trajectories of their arms imply that people factor into their motion plans the effects of gravity and inertia. Several studies have also shown that expectations about the dynamic characteristics of the environment generalize to completely new movements (Conditt *et al.* 1997; Reinkensmeyer *et al.* 2004); such results are important because they suggest that the representation of the dynamics of the environment is not tied to a particular motor motion, but rather appears to exist as a central representation (Milner *et al.* 2007).

Other research suggests that people are able to keep track of the dynamics associated with multiple objects in the same environment on the basis of visual or other sensory clues (Davidson and Wolpert 2004; Reinkensmeyer *et al.* 2004). This ability is implied by the ease with which people are able to switch from manipulating one object to another when performing everyday tasks (Milner *et al.* 2007). For example, raking leaves involves lifting up and pushing around leaves, branches, rakes, bags of leaves, wheelbarrows, etc. Not only are people able to represent the dynamics of individual objects, they are also able to combine them. It has been shown, for example, that people are able to anticipate the force needed to lift two objects together on the basis of having learned the force needed to lift each object individually (Davidson and Wolpert 2004).

Further evidence for the independent representation of dynamic properties comes from research on the use of mental simulation in problem solving (Schwartz 1999; Schwartz and Black 1996). In Schwartz and Black (1996) participants were presented with two glasses of the same height, one narrow and the other wide. Each glass had a line on it to indicate a particular level of imaginary water. The level was at the same height for both glasses. In one condition, people were explicitly asked whether the water in the two glasses would pour out at the same or at different angles. In the second condition, people were asked to close their eyes and to tilt each glass until the point where they believed the water would reach the rim. The correct answer is that as the glasses are tilted, the water will pour out of the wide glass before it pours out of the narrow glass. In the explicit condition, people gave the correct answer only 15% of the time, whereas in the mental simulation condition, people were correct nearly 100% of the time. Amazingly, the process of mentally simulating the event led to more accurate knowledge of the world than predictions based on explicit prior beliefs. In Schwartz (1999), the task was modified slightly to address the question of whether people's simulations were based on kinematics or dynamics. In one condition, the participants were asked to imagine the glass was filled with water while in the other, they were asked to imagine the glass

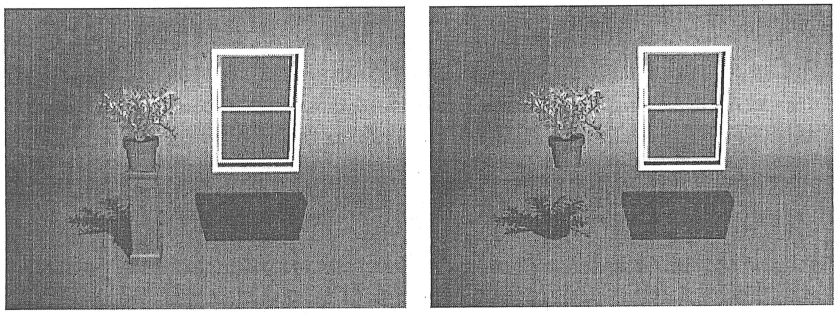

Figure 11.1 Scene adapted from Freyd, Pantzer, and Cheng (1988) in which participants were asked to indicate whether the plant was located in the "same" position once a source of support was removed

was filled with molasses. If people's mental simulations of these events were based purely on kinematics, that is, purely in terms of the geometry of the glass and the liquids, their judgments about the angle at which the liquids should reach the rim should not differ. On the other hand, if their mental simulations of these events were based on dynamics, people should tilt the molasses glass farther than the water glass, since a greater angle is needed (in comparable amounts of time) to make the molasses move to the edge of the rim than is required for water. This is exactly what Schwartz found: people tilted the glass roughly 15° more in the molasses condition than in the water condition. Altogether, the results suggest that people are able to differentiate two events on the basis of dynamic quantities, supporting the view that dynamic quantities such as force can be represented independently of causation.

One last piece of evidence for the independence of force and causation comes from a study in which no obvious causation was present, and yet people inferred the existence of forces. In this study, conducted by Freyd, Pantzer, and Cheng (1988), participants were presented with a scene depicting a potted plant sitting on a pedestal and positioned next to a window, which served as a point of reference (see fig. 11.1). The scene was then replaced by another scene in which the pedestal was removed, but the plant was in exactly the same position as in the earlier scene. This second scene was then replaced with a third scene in which the plant's position was shifted slightly (higher or lower) or remained exactly where it had been before. Freyd *et al.* (1988) reasoned that if people viewed the pedestal as exerting a force on the pot, then people might (implicitly) expect the plant to move downward due to the influence of gravity. The participants' task was to indicate whether the plant in the third display was in the same position as in the second. As predicted, participants were far more likely to report "same" to a shift in the downward than in the upward direction,

supporting the hypothesis that they viewed the pedestal as exerting a force on the potted plant, even in the absence of causation.

In addition to behavioral studies, various patterns in language support the proposal that causation and force can be represented independently of each other. One such pattern is the existence of two relatively large classes of transitive verbs. Certain transitive verbs, lexical causatives, entail the occurrence of a causing and resulting event (e.g., *bend*, *break*, *open*, *drain*, *melt*, *sink*, *bounce*, *roll*, *turn*). Another class of transitive verbs, two-argument activities, also denote actions directed towards an entity (e.g., *clobber*, *hammer*, *hit*, *jab*, *nudge*, *slam*, *shove*, *stroke*, *touch*, *wipe*) but, unlike lexical causatives, they do not strictly entail the occurrence of a change of state of location (Levin and Rappaport Hovav 1994; Pinker 1989; Shibatani 1976b; Song 1996). The difference in meaning between these two classes is revealed when their (possible) results are explicitly denied (Shibatani 1976b). Whereas the possible result to a two-argument activity verb can be explicitly denied (e.g., *John kicked the ice, but nothing happened to it*), the possible result of a lexical causative verb cannot be denied without contradiction (e.g., **John melted the ice, but nothing happened to it*). This difference in semantic entailment is often taken as evidence that that lexical causatives encode for causation whereas two-argument activities do not (Shibatani 1976b; Wolff 2003). Nevertheless, the two verb classes share an important semantic component. As described by Levin (2007), the direct object of both types of verbs can be construed as instantiating the semantic role of *force recipient*, in other words, the target of a transmitted force. According to this analysis, two-argument activities encode for the transmission of force without coding for causation. This analysis is supported by intuition. When we say *He hit the wall*, this does not mean that something was caused; the wall likely remained the same as it was prior to the hitting. However, the sentence does imply that force was imparted since its meaning is quite different from the meaning of the sentence *He touched the wall*, another situation where nothing is caused.

In sum, evidence from both behavioral studies and linguistic analysis indicates that people need not rely on an abstract notion of causation in order to understand the notion of force. Forces can be perceived directly through our senses and appear to be conceptually separate from each other. As stated earlier, Hume's criticism that definitions of causation based on force are circular can be safely rejected.

5 Why causation by omission represents a problem for process theories, and a possible solution

A second major criticism that has been raised against process theories concerns their ability to address the phenomenon of causation by omission. Here we first describe why this phenomenon is a problem for process theories. We then describe a new theory of causation that explains how causation by omission can be handled in terms of causal processes. Lastly, we review empirical evidence in support of this theory.

The main criterion for causation in process theories is the transfer of energy or force. This assumption is clearly at odds with causation by omission. For example, when we say *Lack of rainfall caused the drought*, the cause in this claim, *Lack of rainfall*, is an absence, and obviously no force can be transmitted from an absence. The problem posed by causation by omission has led some philosophers to suggest that there may be two kinds of causation: productive, or positive, causation, in which there is a transfer of force or energy from the cause to the effect, and make-a-difference, or negative, causation, which preserves counterfactual dependencies but not the property of transitivity (Hall 2004; Godfrey-Smith forthcoming; Menzies 2004). Other philosophers have argued that causation by omission is not "really" causation (Beebee 2004; Dowe 2001). For example, Dowe (2001) views causation by omission as "quasi" causation because it does not involve an exchange of conserved quantities. In order to account for statements of causation by omission, Dowe (2001) adopts theoretical machinery from outcome theories, namely, counterfactuals. For other theorists, the inability of process theories to account for causation by omission indicates that such theories are fundamentally flawed (Schaffer 2000; Schulz *et al.* 2007; Woodward 2006). In contrast to this conclusion, we show causation by omission can be specified in terms of causal processing.

To account for causation by omission, we will adopt a proposal implied by several philosophers, namely that causation by omission can be handled in terms of *double prevention* (McGrath 2003; Foot 1967; McMahan 1993). To illustrate double prevention, imagine a situation in which a car is held up off the ground by a jack. A man pushes the jack aside, and the car falls to the ground. This scenario could be described as *The man caused the car to fall to the ground* (even though the force that causes the car to fall to the ground does not come from the force exerted by the man but rather from gravity). What transpires in this example is a sequence of PREVENT relations, or double prevention. First, the jack prevents the car from falling (due to gravity), and then the man prevents the jack from preventing the car from hitting the ground. Critically, instances of double prevention can be paraphrased as instances of causation by omission. As noted earlier, we can say *the man caused the car to fall to the ground*. Alternatively, instead of explicitly naming the man, we can refer to the prevention the man initiated: *lack of a jack caused the car to fall to the ground*. The way a double prevention is described will depend on whether people wish to focus on the entity that prevented the prevention or the absence caused by the prevention. The same idea can be illustrated with the example used earlier: *Lack of rainfall caused the drought*. According to our hypothesis, claims of causation by omission imply double preventions. In this particular example, the implicit double prevention may have been that, for example, *Climate change prevents rainfall* and *Rainfall prevents drought*. Notice that this double prevention licenses the conclusion *Climate change causes drought*, or we could once again choose to highlight what was prevented by climate change and say *Lack of rainfall causes drought*.

Once causation by omission is analyzed as double prevention, it is possible to specify how it might be represented in terms of forces, and hence by a process theory of causation, in particular, the force theory of causation. Our description of the force theory will have five main parts. First, we will introduce the force theory, including some of its assumptions. Second, we will describe how the theory represents individual causal relationships. Third, we will describe how the theory accounts for the joining of causal relations that allow for the generation of overarching relations, or *relation composition*. For example, in order to understand the meaning of a double prevention, two prevention relations must first be joined, which then results in a new causal relation. In a fourth section, we will explain how relation composition can sometimes lead to more than one conclusion, and how the theory explains the relative proportion of possible conclusions. Finally, we will describe how the theory accounts for the representation of causation by omission and causation of an absence in detail in terms of the relation composition of PREVENT relations.

6 Force theory

According to the force theory, people specify causal relations in terms of configurations of forces that can be represented as vector quantities that can exert an influence. The forces may be physical, psychological (e.g., intentions), or social (e.g., peer pressure) (Wolff 2007). The force theory is primarily described at the algorithmic level; in other words, it is meant to account for the actual cognitive operations that people perform when they reason. However, certain aspects of the model are described at the computational level, that is, at a level that is intended to simply predict human performance while adopting computational procedures that are not psychologically plausible. We will explicitly note those parts of the theory that extend beyond the algorithmic to the computational level.

We assume that the specification of forces must be partially symbolic, especially when the forces involved are psychological or social. They must also be symbolic in that the magnitudes of the vectors used in the representation of causal relations may be somewhat arbitrary (Wolff 2007). At the same time, it is also assumed that reasoning with forces is partially iconic. According to the force theory, vector representations in the mind support the same kinds of processes that occur among forces in the world (e.g., vector addition, subtraction); thus, according to the theory, specifying causal relations involves a partial "re-enactment" of how forces interact in the world.

The process of re-enactment in the mind begins with a specification of the quantities that produce causal relations, namely forces. A key test of the claim that the force theory specifies the quantities that produce causal relations is whether the representations specified by the model can be entered into a physics simulator to produce animations reflecting real-world events and whether the animations are recognized by people as causal in the ways specified by the theory. Such a test is described below. In being able to meet this test, the force theory provides an account of the representations that might drive the analog mental simulations that people "watch" in their minds to recognize and identify causal relations (see Hegarty 2004).

In the force theory, uncertainty is built into the representation of causation because of uncertainty about the magnitudes of the vectors in a configuration of forces. For individual causal relations, this uncertainty will be of little or no consequence: while people may not know the precise magnitudes of the vectors involved, they can ascribe relative magnitudes, which is enough to classify a configuration of forces as causal, preventative, or allowing. However, as discussed below, when configurations of force are added together to form networks or chains of configurations, probabilistic outcomes will emerge from deterministic processes. Further important assumptions of the force theory are discussed in Wolff (2007).

6.1 *Representing individual causal relations in the force theory*

The force theory extends Wolff's (2007) dynamics model of causation and Talmy's (1988) theory of force dynamics in specifying how individual causal relations might be represented in configurations of force. According to the dynamics model, the concept of CAUSE and related concepts involve interactions between two main entities: an *affector* and a *patient* (the entity acted on by the affector). It holds that the different kinds of causal relationships can be specified in terms of three dimensions: (a) the *tendency* of the patient for an endstate, (b) the presence or absence of *concordance* between the affector and the patient, and (c) *progress toward the endstate* (essentially, whether the result occurs). Table 11.1 summarizes how these dimensions differentiate the concepts of CAUSE, HELP/ENABLE/ALLOW, and PREVENT. For example, according to the dynamics model, when we say *High winds caused the tree to fall down*, we mean that the patient (the tree) had no tendency to fall (Tendency = No), the affector (the wind) acted against the patient (Concordance = No) and the result (falling down) occurred (Endstate approached = Yes).

Table 11.1 *Representations of several causal concepts*

	Patient tendency for endstate	Affector–patient concordance	Endstate approached
CAUSE	No	No	Yes
HELP/ENABLE/ALLOW	Yes	Yes	Yes
PREVENT	Yes	No	No

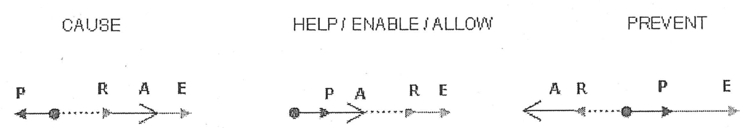

Figure 11.2 Configurations of forces associated with CAUSE, HELP/ENABLE/ALLOW, and PREVENT; **A** = the affector force, **P** = the patient force, **R** = the resultant force; **E** = endstate vector, which is a position vector, not a force

The dynamics model specifies how these three dimensions are captured in terms of configurations of force vectors. Sample configurations of forces for CAUSE, HELP/ ENABLE /ALLOW, and PREVENT are depicted in fig. 11.2. As is customary, the free-body diagrams in fig. 11.2 show forces acting on only one object, in these cases, the patient entity. They do not show the location of the affector entity, only the direction and magnitude of the affector's force on the patient (i.e. **A**). Similarly, they do not show the location of the endstate, just the vector that points from the patient to the endstate (i.e. **E**). In each of the configurations shown in fig. 11.2, the patient entity is also associated with a force (i.e. **P**). If the patient were a boat, the patient force might correspond to the force generated by the boat's motor; if the patient were a rock on the ground, the patient force might correspond to the force of friction (or the tendency to resist movement in a particular direction due to frictional forces). When the patient has a tendency for the endstate, the patient vector, **P**, will point in the same direction as the endstate vector, **E**; when the patient does not have a tendency for the endstate, the patient vector will point in a different direction from **E**. When the patient and the affector are in concordance, their respective vectors will point in the same direction; otherwise, they will point in different directions. Finally, the patient entity will approach the endstate when the resultant of the **A** and **P** vectors, **R**, is in the same direction as the endstate vector, **E**.

Support for the dynamics model's account of CAUSE, HELP/ENABLE/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 (Wolff 2007). (The animations can be viewed at http://userwww.service.emory.edu/~pwolff/CLSAnimations.htm.) In these experiments, the very same physical forces used to generate physical scenes were used as inputs into a computer model to predict how those scenes would be described. As reported in Wolff (2007 and Wolff and Zettergren 2002), the fit between the predictions of the model and people's descriptions of the events was strong.

6.2 *Combining relations in the force theory*

Whereas the dynamics model is an account of how people represent individual relations, the force theory is an account of how people generate new relations by combining them, a process referred to in mathematics as relation composition. For example, given that *nerve damage causes pain* and *pain causes lost work-days*, an overarching causal relation can be generated via the composition of the two cause relations to produce *nerve damage causes lost workdays*. In the following sections we show how these two parts of relation composition – the joining of relations and the generation of summary conclusions – are accomplished.

In the force theory, the mechanism for combining relations depends on whether the initial relation in a pair of relations is generative or preventative. If the initial relation is generative (CAUSE, HELP or ALLOW), then the first relation is connected to the second relation by using the resultant of the first relation as the affector in the second relation. For example, as shown in fig. 11.3, if the causal chain involves a sequence of CAUSE relations, the resultant in the first premise (**BA**) is the affector in the second premise ($\mathbf{B_{BA}}$). This sequence can be exemplified by a chain of three marbles, A, B and C in which A first hits B, which in turn hits C. The force that sends B into motion is based on the resultant of the force acting on A moving towards B, minus the force slowing it down (i.e. friction); in other words, the force acting on B would be based on the resultant of the forces acting on A. The force acting on C would, in turn, be based on the resultant of the forces acting on B.

When the initial relation in a pair of relations is preventative (PREVENT), the process of relation composition proceeds in a different manner. Note that if A first prevents B, B cannot act on C because B has been prevented. The way such chains are understood, then, is that an interaction first occurs between B and C, and then A acts on B. In terms of forces, as shown in fig. 11.3, when a PREVENT relation is followed by another relation (CAUSE, ALLOW, or PREVENT), the resultant of the second premise (**CB**) is the patient vector in the first premise ($\mathbf{B_{CB}}$). The intuition behind this way of connecting the premises can be illustrated with a real-world example of double prevention: pulling a plug to allow water to flow down the drain. This sequence of prevents begins

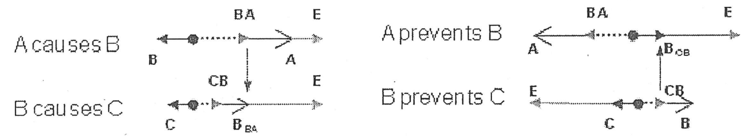

Figure 11.3 On the left side, two CAUSE relations are combined using the resultant force from the first cause relation (**BA**) as the affector force in the second cause relation ($\mathbf{B_{BA}}$). On the right side, a PREVENT relation is combined with another PREVENT relation using the resultant of the PREVENT relation in the second premise as the patient vector in the PREVENT relation in the first premise

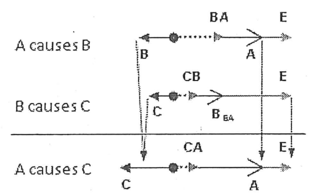

Figure 11.4 The affector force in the conclusion, **A**, is the affector force in the first relation, **A**. The endstate in the conclusion is the endstate vector from the last premise. The patient force in the conclusion, **C**, is based on the vector addition of the patient forces, **B** and **C** in the premises

with the plug (B) preventing the water (C), that is, with the second premise in a double prevention. With this prevention in place, the next step is for someone (A) to pull the plug (B), which constitutes the first prevent relation. Thus, in double prevention, the order of causation, in a sense, occurs in reverse order of the premises. In the theory, this reversal is captured by using the resultant force in the second premise as the patient vector in the first premise.

6.3 *Generating a conclusion*

Regardless of the type of causal chain, the manner in which an overall conclusion is reached is the same. As depicted in fig. 11.4, the affector in the conclusion is the affector from the first premise; the endstate in the conclusion is the endstate from the last premise; and the patient in the conclusion is the resultant of the patient vectors in the premises. Intuitively, the patient vector specifies what would happen in the absence of the affector force since it is based on all of the forces in the chain except the affector force. The patient force in the conclusion allows people to evaluate the truth value of the counterfactual *if a were not present, b would not occur*. According to some theorists, evaluating such a counterfactual is a key part of establishing a causal relation (see Lewis 1973, 2000; Mackie 1974; see also Mandel and Lehman 1996; Spellman and Mandel 1999; Spellman, Kincannon, and Stose 2005). The force theory specifies the knowledge that allows such counterfactuals to be evaluated.

6.4 *Accounting for multiple conclusions*

Prior research indicates that people's representations of forces are underspecified with respect to magnitude (Wolff 2007). Not knowing the exact magnitude of the forces adds indeterminacy to people's representations of causation. The effects of this indeterminacy also emerge when configurations of force are combined: variations in the magnitudes of the forces can sometimes lead to more than one summary conclusion or to a conclusion of reduced strength due to the presence of un-categorizable summary configurations. One type of relation composition that can lead to multiple conclusions is that of double prevention. As argued by several researchers, the composition of two PREVENT relations can sometimes lead to a CAUSE relation and other times to an ALLOW relation (McGrath 2003; Barbey and Wolff 2006, 2007; Sloman *et al.* 2009; Chaigneau and Barbey 2008). The intuition behind this claim can be illustrated with everyday instances of double prevention. Imagine, for example, causing/allowing a pencil to fall to the floor by letting go of it. Initially, there is a PREVENT relationship between your hand and the pencil: your hand "prevents" the pencil from falling to the floor. With this pre-condition in place, you open your hand, thereby preventing the prevention. The magnitudes of the forces are unclear in this example and, as a consequence, the double prevention is open to either a CAUSE or ALLOW interpretation: we can say either *I allowed the pencil to fall to the floor* or *I caused the pencil to fall to the floor.*

The force theory predicts that double preventions can lead to either CAUSE or ALLOW conclusions, depending on the magnitude of the patient vectors in the premises. As shown on the left side of fig. 11.5, double preventions lead to CAUSE interpretations when the $\mathbf{B_{CB}}$ patient vector is greater in magnitude than the **C** patient vector; as a consequence, the patient vector in the conclusion points away from the endstate when these two vectors are added together. When the relative magnitude of these two vectors is reversed, as in the right side of fig. 11.5, the conclusion is ALLOW.

Given that a sequence of PREVENT relations can lead to either a CAUSE or an ALLOW conclusion, it raises the more general question of which conclusion is more likely. One way to find out is to systematically vary the magnitudes of the vectors and then tally the number of conclusions that follow from the different combinations of magnitudes. A program has been written that conducts such a simulation process (http://userwww.service.emory.edu/~pwolff/Transitivedynamics.htm). One part of the program allows users to create chains

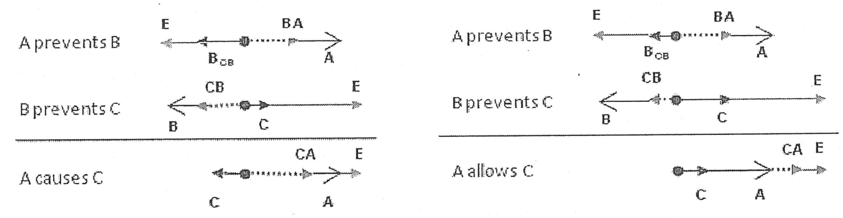

Figure 11.5 The composition of two PREVENT relations can either lead to a CAUSE or ALLOW conclusion

of any length and of any combination of relations by manipulating the magnitudes of vectors with the mouse. A second part of the program implements a simulation process that exhaustively varies the magnitudes of the vectors (assuming a uniform distribution) within the constraints set by the relations in the premises. The program then tallies the percentages of conclusions that are generated from all of the possible combinations of magnitudes. Using this procedure, the program finds, for example, that the conclusions associated with double prevention are 62% ALLOW and 38% CAUSE. A second way to determine the proportions of different conclusions is to use integral calculus, as described in Barbey and Wolff (ms.).

6.5 *Representing ALLOW/ENABLE*

We propose that the concepts of ALLOW and ENABLE are complex relations derived from the composition of two PREVENT relations. In other words, when the composition of two PREVENT relations results in a conclusion in which the affector and patient vectors both point toward the endstate, the resulting conclusion is interpreted as either ALLOW or ENABLE. The idea that ALLOW and ENABLE are based on a series of PREVENT relations is consistent with prior work in philosophy (McGrath 2003; Foot 1967; McMahan 1993), psychology (Barbey and Wolff 2006, 2007; Sloman *et al.* 2009; Chaigneau and Barbey 2008; Wolff 2007), and linguistics (Talmy 1988). One of the benefits of defining ALLOW/ENABLE in terms of double prevention is that it accounts for the intuition that the affector in an ALLOW relation is a necessary condition for the result (Reinhart 2002; Goldvarg and Johnson-Laird 2001). In generating the predictions for different kinds of causal compositions, ALLOW will be treated as double preventions.

7 Explaining causation by omission

As discussed earlier, the way the force theory handles causation by omission is in terms of double prevention. Consider, again, the example of double prevention in which a person pulls a plug and allows water to flow down the drain. Instead of focusing on the entity that prevented the prevention, the same situation could be described in terms of the absence that was created, specifically, *Absence of a plug allowed/caused water to flow down the drain*. Similarly, in the pencil scenario, we could say *Lack of support allowed/caused the pencil to fall to the floor*. As discussed earlier, how a double prevention is described will depend on whether the speaker wishes to focus on the entity/event that prevented the prevention or the absence caused by the prevention.

As noted earlier, once causation by omission is analyzed as double prevention, it becomes possible to specify how it might be represented in terms of forces and, hence, by a process theory of causation. For example, consider the chain of forces and frames from an animation generated from the forces in fig. 11.6. The chain of forces at the top of fig. 11.6 instantiates double prevention. Relevant information about these forces was entered into a physics simulator to produce the animation depicted in fig. 11.6. In the beginning of the animation (left panel), car C approaches the line. Car B then approaches car C and prevents it from crossing the line (middle panel). Car A then pulls car B away, preventing the prevention (middle panel). Finally, with car B out of the way, car C crosses the line (right panel). As with the examples above, this sequence of prevents leads to a positive outcome. However, in this case, it seems more natural to say *Car A allowed car C to cross the line* than to say *Car A caused car C to cross the line*.[1]

In any case, the double prevention depicted in fig. 11.6 can be described in two ways: *Car A allowed car C to cross the line* or *The absence of car B allowed car C to cross the line*. Importantly, although car B exerts a force on car C, and car A exerts a force on car B, these two forces are in opposite directions. Because they are in opposite directions, the force car A exerts on car B is not transferred from car B to car C. Rather, the force that car A exerts on car B *removes* the force that car B exerts on car C. Further, note that there is no energy tradeoff between cars A and C. When energy is transferred, the energy gained by one entity is balanced by the energy lost by another. In the animation depicted in fig. 11.6, as car A accelerates, so does car C, indicating that kinetic energy is not transferred from car A to car C. These observations are important because they demonstrate that transmission of force is not a necessary condition of causation; causation can also come about from the removal of a force. The idea that causation need not involve a transmission of force conflicts with a major assumption of previous process theories. The force

[1] In the literature on double prevention, it has been noted that the sequence resulting from a double prevention sometimes seems causal, while at other times it does not (Hall 2004; Livengood and Machery 2007). In those cases where it sounds odd to describe a double prevention in terms of "cause," an "allow" paraphrase is often acceptable. Notice that *Car A allowed car C to cross the line* can also be described as *The absence of car B allowed car C to cross the line*.

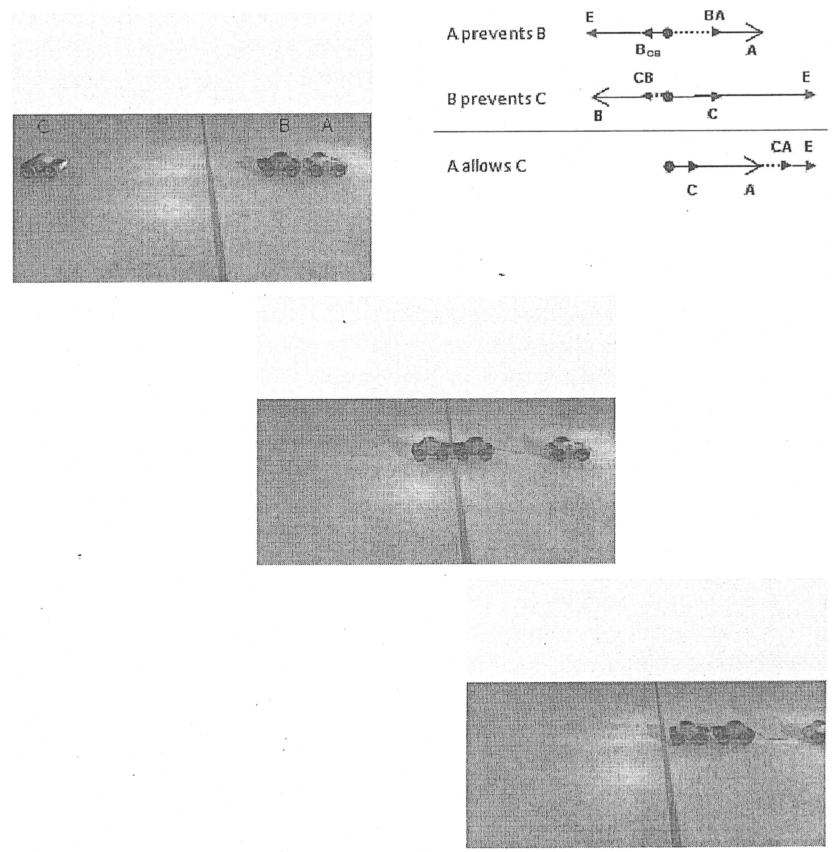

Figure 11.6 The configuration of forces in the top panel, which depicts a PREVENT ∘ PREVENT composition, was entered into a physics simulator to produce the movements of the cars in the animation depicted in the still frames in the bottom panel. First, car C attempts to cross the line but is prevented by car B, which approaches car C. Then, car A pulls car B away from car C with a rope, preventing car B from preventing car C. Finally, with car B out of the way, car C crosses the line

theory explains how this assumption can be let go by process theories while maintaining the commitment to the centrality of energy and force in people's representations of causation. As discussed earlier, the reason why causation by omission has been a problem for process theories is because these theories held that causation was defined in terms of transmission. Removing this assumption makes it possible for process theories to account for causation by omission as well as several other phenomena, including the meaning of ALLOW and the causation of absences (see Barbey and Wolff ms.).

In the following section, we describe two lines of support for the force theory and its account of causation by omission. In one line, we show that the force theory is able to predict the conclusions people generate for a wide range of causal compositions, including those involving causation by omission. In a second line, we show that the force theory is able to account for the kinds of expressions people choose to use when describing animations of physical events.

8 Evidence in support of the force theory

The force theory is capable of explaining the relational composition of a wide range of causal chains. Some of the predictions of the theory were tested in an experiment reported in Barbey and Wolff (2006; ms.). In that experiment, participants (N = 40) read two-relation causal chains constructed from real-world causal statements found on the internet. For example, for the argument A causes B and B causes C, participants read sentences like *Factories cause pollution*, *Pollution causes global warming*. They also read sentences involving causation by omission. For example, for the argument not-A causes not-B and B causes C, people saw statements like *Leaf loss causes lack of shade* and *Shade causes cooling*. Six real-world instantiations were found for all thirty-two argument types shown in table 11.2 for a total of 192 arguments. In table 11.2, the columns show all possible 1st relations and the rows show the different possible 2nd relations. After reading the causal chain, participants chose the relation between the unconnected A and C terms that sounded most natural from a list of ten possible conclusions (*A causes C*, *A allows C*, *A prevents C*, *A causes not-C*, *A allows not-C*, *A prevents not-C*, *not-A causes C*, *not-A allows C*, *not-A prevents C*, or *none of the above*). In the actual experiment, the A, B, and C terms were filled in with the terms named in the causal chain.

The results in table 11.2 show that participants' relational compositions were well explained by the force theory. As a measure of fit, the average Pearson correlation between the mean response to each item and the force theory was .86, a correlation that was significantly different from chance, $p < .0001$. According to the force theory, double prevents should lead to ALLOW responses 62% of the time and CAUSE responses 38% of the time. As shown in table 11.3, the result supported the prediction. The most frequent response to double preventions was ALLOW, and ALLOW responses were chosen more often than CAUSE responses, 39% to 8%. As shown in table 11.2, fourteen of the thirty-two causal chains involved causation by omission. The force theory correctly predicted the most frequent response for twelve of these chains, a result that was significantly greater than chance by a binomial test, $p = 0.013$. The results support the hypothesis that the force theory describes the processes that people use to compose causal relations, including causal compositions involving causation by omission.

Table 11.2 *Results from Experiment 1 of Barbey and Wolff (ms.) showing the predicted (in bold) and observed percentages of response types*

	1st relation			
2nd relation	A causes B	A allows B	A prevents B	¬A causes B
B causes C	**C (100%)** C (78%)	**A (76%) C (24%)** A (66%) C (23%)	**P (37%)** P (93%)	**¬C (50%) P (43%)** ¬C (46%) P (26%)
B allows C	**A (76) C (24%)** A (69) C (15%)	**A (100%)** A (89%)	**P (30%)** P (82%)	**¬A (90%) ¬C (9%)** ¬A (28%) ¬C (13%)
B prevents C	**P (100%)** P (77%)	**P (22%)** P (69%)	**A (62%) C (38%)** A (39%) C (8%)	**¬P (51%)** ¬P (38%)
¬B causes C	**P (59%)** P (62%)	**P (5%)** P (71%)	**C (49%) A (22%)** C (58%) A (18%)	**¬P (20%)** C (23%) ¬P (20%)
	A causes ¬B	A allows ¬B	A prevents ¬B	¬A causes ¬B
B causes C	**P (59%)** P (59%)	**P (5%)** P (61%)	**C (49%) A (22%)** C (39%) A (20%)	**¬P (19%)** ¬P (34%)
B allows C	**P (17%)** P (70%)	**P (69%)** P (68%)	**A (92%) C (7%)** A (53%) C (8%)	**¬P (47%)*** A (43%) ¬P (29%)
B prevents C	**C (61%) A (39%)** C (60%) A (28%)	**A (95%) C (5%)** A (50%) C (18%)	**P (49%)** P (65%)	**¬A (75%)* ¬C (23%)** P (37%) ¬C (30%) ¬A (24%)
¬B causes C	**C (54%) A (13%)** C (82%) A (7%)	**A (62%) C (35%)** A (60%) C (26%)	**P (32%)** P (69%)	**¬C (47%) ¬A (31%)** ¬C (34%) P (30%) ¬A (3%)

Note: C = Cause, A = Allow, P = Prevent, ¬C = Not A causes B, ¬A = Not A allows B, ¬P = Not A prevents B; * = missed prediction

The results described above show that the force theory is consistent with the way that people understand and express causal relations. However, one of the most important features of the force theory is that it offers an account of how people might be able to recover causal relations from the world on the basis of a single observation. This prediction was examined in detail – and supported – for single relations in both physical and social domains in Wolff (2007). The force theory also predicts that it should be possible to interpret causal chains on the basis of a single observation. In addition, the theory explains how the phenomenon of double prevention – and by extension, the concept of ALLOW and causation by omission – can be instantiated in physical processes. These predictions were tested in the following experiment.

Four 3D animations were made from an animation package called 3D Studio Max (ver. 8). Each animation involved three cars, labeled A, B, or C, acting on each other by pushing and pulling, and a line on the ground, which defined an endstate. The four types of animations were a CAUSE/CAUSE chain, a

Table 11.3 *Percentage of responses for four types of causal chains*

	Chain type			
	CAUSE/CAUSE	CAUSE/PREVENT	P/P-CAUSE	P/P-ALLOW
A caused C	**90%**	–	**47%**	10%
A allowed C	6.7%	–	**43%**	**70%**
A prevented C	–	**90%**	–	–
None of the above	3%	–	10%	20%
B caused C	**63%**	–	7%	7%
B allowed C	27%	–	17%	27%
B prevented C	–	**87%**	13%	17%
None of the above	10%	13%	**63%**	**50%**
Not A caused C	3%	–	7%	–
Not A allowed C	7%	7%	7%	13%
Not A prevented C	3%	13%	17%	13%
None of the above	**87%**	**80%**	**70%**	**73%**
Not B caused C	23%	3%	7%	7%
Not B allowed C	13%	10%	**50%**	**50%**
Not B prevented C	10%	23%	3%	–
None of the above	**52%**	**56%**	40%	43%

CAUSE/PREVENT chain, and two kinds of PREVENT/PREVENT chains. The direction and magnitude of the cars were calculated using a physical simulator called Havok Reactor. The input into the physics simulator consisted of forces generated by the computer version of the force theory. The average animation lasted 8 seconds. Each time an animation was presented, participants (N = 30) were presented with four possible descriptions. In condition 1, participants chose the best description of the animation from a list of four options: (a) *A caused C to cross the line*, (b) *A allowed C to cross the line*, (c) *A prevented C from crossing the line*, and (d) *None of the sentences above are applicable to the scene*. The options in condition 2 were the same as those in condition 1 except that all of the As were replaced with Bs; for example, for the CAUSE option, participants were given the sentence *B caused C to cross the line*. In condition 3, the sentences were also the same as in condition 1, except that the cause was described in terms of its absence; for example, for the CAUSE option, participants were given the sentence *The absence of A's influence caused C to cross the line*. In condition 4, the sentences were the same as in condition 3, except that the As were replaced with Bs; hence, in this condition, the CAUSE option was *The absence of B's influence caused C to cross the line*. The four conditions were run within participants, in other words, participants saw each animation four times and each time with a different set of options.

In the case of the CAUSE/CAUSE chain (C/C), car A hit car B, which then hit car C, pushing it over the line. For this animation, we predicted that people would choose the description *Car A caused car C to cross the line*. In the case of the CAUSE/PREVENT chain (C/P), car A hit car B, which then hit car C, blocking it from crossing over the line. For this animation, we predicted that people would choose the description *Car A prevented car C from crossing the line*. As stated above, there were two PREVENT/PREVENT chains. For the first of these chains, P/P-CAUSE, the relative magnitudes were such that they implied a CAUSE response (see left side of fig. 11.5). For the second PREVENT/PREVENT chain, P/P-ALLOW, the relative magnitude of the forces was such that they implied an ALLOW response (see right side of fig. 11.5). The P/P-ALLOW animation was the same as the one depicted in the bottom of fig. 11.6.

The key predictions with regard to the two PREVENT/PREVENT chains were that the proportion of CAUSE responses would be higher for the P/P-CAUSE chain than for the P/P-ALLOW chain, and that the proportion of ALLOW responses would be higher for the P/P-ALLOW chain than for the P/P-CAUSE chain. Finally, for the PREVENT/PREVENT chains, but not the other chains, we predicted that people would be willing to describe the causation in terms of absences, specifically, that for the two P/P chains, they would be willing to say *The absence of B's influence caused/allowed C to cross the line*.

The results in table 11.3 show that these predictions were supported. For the C/C chain, people were very happy to say that *Car A caused car C to cross the line* as well as *Car B caused car C to cross the line*. For the C/P chain, participants indicated that *Car A prevented car C from crossing the line*, and that *Car B prevented C from crossing the line*. Of primary interest was how people described the causal chains implementing double prevention. As predicted, people preferred to describe the relationship between the A and C cars in the P/P-ALLOW chain with an ALLOW description, but they were also willing to use a CAUSE description. For the P/P-CAUSE chain, people split in their preference between CAUSE and ALLOW descriptions; importantly, though, the number of CAUSE descriptions for P/P-CAUSE chains was higher than for the P/P-ALLOW chains.

Of particular interest, people were willing to describe the P/P chains – but not others – in terms of the absence of an influence. For both types of P/P chains, participants were happy to say *The absence of B's influence allowed C to cross the line*. The results support the hypothesis that people conceptualize causation by omission in terms of double prevention.

9 Summary

In this chapter we contrasted two general approaches to causation. We argued that outcome theories define causation in terms of the outward signs of causal relationships while process approaches define causation in terms of the processes that produce causation in the world. We argued that the quantities that are essential to causation in the world are also central to causation in the mind. In support of this position, we discussed how several of the most prominent characteristics of causation – spatial and temporal contiguity, mechanism, and temporal priority – fall out naturally from representations of causation that are based on dynamics, but do not fall out of representations that specify the outward signs of causation, loosely speaking, their kinematics. In further support of process theories, we reviewed literature suggesting that people are able to represent the dynamic properties of the environment. Finally, we addressed the most serious challenges that have been raised against process approaches to causation, namely Hume's criticism that theories based on force are circular and the problem of how to represent causation by omission. We showed that these challenges can be handled by thinking about causation in terms of the configurations of forces. The key insight is that causation is based not only on the transmission of force, but also on its removal. In support of this proposal, we described the results of two experiments. In one, we showed that a particular process model, the force theory, was able to predict the relational compositions associated with linguistic descriptions. In a second experiment, we demonstrated the ability of the force theory to predict the composition of relations from visually presented stimuli and, in particular, its ability to differentiate double preventions that are typically construed as CAUSE relations from those that are typically construed as ALLOW relations. In addition, we showed that both types of these double prevention chains can be described in terms of the absence of a force.

10 Conclusions

Much of the recent research on causal cognition has been dominated by outcome theories of causation, that is, the proposal that causal knowledge is specified in terms of actual or possible outcomes. The emphasis is understandable since outcome theories have been able to address in a rigorous fashion some of the most interesting aspects of causal cognition, in particular, the problems of how people put causal relationships together into larger structures and then reason over those structures. The force theory represents the first process theory to address these more demanding problems. In the force theory, we know the relevant entities (patient, affector, endstate), their assignment to representational elements of causal knowledge (configurations of force), their meaning (quantitative force vectors), and how their structure maps (sometimes directly) onto the structure of causal events in the world, demonstrating for the first time how our understanding of forces in the world might provide the basis for our ability to represent and reason about causal relations.

References

Ahn, W., and Bailenson, J. (1996). Mechanism-based explanations of causal attribution: An explanation of conjunction and discounting effect. Cognitive Psychology 31: 82-123.

Ahn, W., and Kalish, C. W. (2000). The role of mechanism beliefs in causal reasoning. In F. C. Keil and R. A. Wilson (Eds.), *Explanation and cognition* (pp. 199-225). Cambridge, MA: The MIT Press.

Ahn, W., Kalish, C. W., Medin, D. L., and Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. Cognition 54: 299-352.

Aronson, J. L. (1971). On the grammar of 'CAUSE'. *Synthese, 22*, 414-430.

Barbey, A. K., and Wolff, P. (2006). Causal reasoning from forces. Proceedings of the 28thAnnual Conference of the Cognitive Science Society (p. 2439). Mahwah, NJ: Lawrence Erlbaum.

Barbey, A. K. and Wolff, P. (2007). Learning causal structure from reasoning. Proceedings of the 29th Annual Conference of the Cognitive Science Society (pp. 713-718). Mahwah, NJ: Lawrence Erlbaum.

Barbey, A. K. and Wolff, P. (submitted). Composing causal relations in force dynamics.

Beebee, H. (2004). Causing and Nothingness. In, J. Collins, N. Hall, and L. Paul, (ed.), Causation and Counterfactuals (291-308). Cambridge, MA: MIT Press.

Bullock, M., Gelman, R., and Baillargeon, R. (1982). The development of causal reasoning. In W.Friedman (Ed.), The developmental psychology of time (pp. 209-255). London: Academic Press.

Chaigneau, S. E., and Barbey, A. K. (2008). Assessing psychological theories of causal meaning and inference. In B. C. Love, K. McRae, and V. M. Sloutsky (Eds.), Proceedings of the 30th Annual Conference of the Cognitive Science Society (pp. 1111-1116). Austin, TX: Cognitive Science Society.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367-405.

Cheng, P. W., and Novick, L. R. (1991). Causes versus enabling conditions. Cognition 40: 83-120.

Cheng, P. W., and Novick, L. R. (1992). Covariation in natural causal induction. Psychological Review 99: 365-382.

Conditt, M. A., Gandolfo, F., and Mussa-Ivaldi, F. A. (1997). The motor system does not learn the dynamics of the arm by rote memorization of past experience. Journal of Neurophysiology 78: 554- 560.

Davidson, P. R., and Wolpert, D. M. (2004). Internal models underlying grasp can be additively combined. Experimental Brain Research 155: 334 – 340.

Dowe, P. (2000). Physical causation. Cambridge: Cambridge University Press.

Dowe, P. (2001). A counterfactual theory of prevention and “causation” by omission. Australasian Journal of Philosophy 79: 216-226.

Dowe, P. (2007). Causal processes. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu.

Fair, D. (1979). Causation and the flow of energy. Erkenntnis 14: 219-250.

Foot, P. (1967). The problem of abortion and the doctrine of double effect. Oxford Review 5, 5-15; reprinted in B. Steinbock and A. Norcross (Eds.), Killing and letting die, 2nd ed. (pp. 266-279). NY: Fordham University Press.

Freyd, J. J., Pantzer, T. M., and Cheng, J. L. (1988). Representing statics as forces in equilibrium. Journal of Experimenal Psychology: General 117: 395 – 407.

Goldvarg, E., and Johnson-Laird, P. (2001). Naive causality: A mental model theory of causal meaning and reasoning. Cognitive Science 25: 565-610.

Godfrey-Smith, P. (in press). Causal pluralism. In H. Beebee, C. Hitchcock, and P. Menzies, Oxford Handbook of Causation.

Gopnik, A., Glymour, C., Sobel, D., Shulz, L., Kushnir, T., and Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. Psychological Review 111: 1-31.

Hall, N. (2004). Two concepts of causation. In J. Collins, N. Hall, and L. Paul (Eds.), Causation and counterfactuals (pp. 225-276). Cambridge, MA: MIT Press.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. TRENDS in Cognitive Sciences 8: 280 –285.

Hinder, M. R., and Milner, T. E. (2003). The case for an internal dynamics model versus equilibrium point control in human movement. Journal of Physiology 549: 953 – 963.

Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. Journal of Philosophy 98: 273-99.

Hitchcock, C. (2002). Probabilistic Causation. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/causation-probabilistic/

Hume, D. (1737/1978). A Treatise of Human Nature, edited by L. A. Selby-Bigge, 2nd edition revised by P. H. Nidditch, Oxford: Oxford University Press.

Imamizu, H., Uno, Y., and Kawato, M. (1995). Internal representations of the motor apparatus:Implications from generalization in visuomotor learning. Journal of Experimental Psychology: Human Perception and Performance 21: 1174 – 1198.

Kawato, M. (1999). Internal models for motor control and trajectory planning. Current opinion in neurobiology 9: 718 – 727.

Kruschke, J. K., and Fragassi, M. M. (1996). The perception of causality: Feature binding in interacting objects. In G. W. Cottrell (Ed.), Proceedings of the Eighteenth Annual conference of the Cognitive Science Society (pp. 441-446). Hillsdale, NJ: Erlbaum.

Kurtzer, I., Herter, T. M., and Scott, S. H. (2005). Random change in cortical load representation suggests distinct control of posture and movement. Nature Neuroscience 8: 498 – 504.

Leslie, A. (1984). Spatiotemporal continuity and the perception of causality in infants. Perception 13: 287-305.

Leslie, A. M., & Keeble S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*, 265-288.

Lesser, H. (1977). The growth of perceived causality in children. The Journal of Genetic Psychology 130: 145-152.

Levin, B., and Rappaport Hovav, M. (1994). A preliminary analysis of causative verbs in English. Lingua 92: 35–77.

Levin, B. (2007). The lexical semantics of verbs I: Introduction. 2007 Summer Institute of Linguistics at Stanford University, Handout.

Lewis, D. (1973). Causation. Journal of Philosophy 70: 556-567.

Lewis, D. (2000). Causation as influence. Journal of Philosophy 97: 182-197.

Livengood, J., & Machery, E. (2007). The folk probably don't think what you think they think: Experiments on causation by absence. *Midwest Studies in Philosophy, XXXI*, 107 – 127.

Mackie, J. L. (1974). The cement of the universe. Oxford: Oxford University Press.

Mandel, D. R., and Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. Journal of Personality and Social Psychology 71: 450–463.

McGrath, S. (2003). Causation and the making/allowing distinction. Philosophical Studies114: 81-106.

McGraph, S. (2005). Causation by omission: A dilemma. Philosophical Studies 123:125 – 148.

McMahan, J. (1993). Killing, letting die, and withdrawing aid. Ethics 103: 250–279.

Menzies, P. (2004). Difference-making in context. In J. Collins, N. Hall, and L. Paul (Eds.), Causation and counterfactuals (pp. 139-180). Cambridge, MA: MIT Press.

Michotte, A. E. (1946/1963). The perception of causality. New York: Basic Books. (Original work published 1946.)

Milner, T. E., and Franklin, D.W. (2005). Impedance control and internal model use during the initial stage of adaptation to novel dynamics in humans. Journal of Physiology 567: 651–664.

Milner, T. E., Franklin, D. W., Imamizu, H., and Kawato, M. (2007). Central control of grasp: Manipulation of objects with complex and simple dynamics. Neuroimage 36: 388 – 395.

Oakes, L. M. (1994). The development of infants' use of continuity cues in their perception of causality. Developmental Psychology 30: 869-879.

Ohta, K., and Laboissière, R. (2006). Underlying principles of trajectory formation for human movement in dynamical environments. International Congress Series 1291: 97 – 100.

Papaxanthis, C., Pozzo, T., and McIntyre, J. (2005). Kinematic and dynamic processes for the control of pointing movements in humans revealed by short-term exposure to microgravity. Neuroscience 135: 371 – 383.

Pinker, S. (1989). Learnability and cognition: the acquisition of argument structure. Cambridge, MA: MIT Press.

Reinhart, T. (2002). The theta system: An overview. Theoretical Linguistics 28: 229-290.

Reinkensmeyer, D. J., Emken, J. L., Crammer, S. C. (2004). Robotics, motor learning, and neurological recovery. Annual Review of Biomedical Engineering 6: 497 – 525.

Russell, B. (1948). Human Knowledge. New York: Simon and Schuster.

Shadmehr R, Mussa-Ivaldi F. A. (1994). Adaptive representation of dynamics during learning of a motor task. Journal of Neuroscience 14: 3208–3224.

Salmon, W. (1994). Causality without counterfactuals. Philosophy of Science 61: 297-312.

Salmon, W. (1998). Causality and explanation. Oxford: Oxford University Press.

Schaffer, J. (2000). Causation by disconnection. Philosophy of Science 67: 285 – 300.

Scholl, B. J., and Tremoulet, P. D. (2000). Perceptual causality and animacy. Trends in Cognitive Sciences 4: 299-309.

Schulz, L., Kushnir, T., and Gopnik, A. (2007). Learning from doing: Intervention and causal inference. In A. Gopnick and L. Schulz (Eds.), Causal learning: Psychology, philosophy, and computation (pp. 67–85). Oxford: Oxford University Press.

Schwartz, D. L. (1999). Physical imagery: Kinematic versus Dynamic Models. Cognitive Psychology 38: 433 – 464.

Schwartz, D. L., and Black, J. B. (1996). Analog imagery in mental model reasoning: Depictive models. Cognitive Psychology 30: 154 – 219.

Shibatani, M. (1976). The grammar of causative constructions: a conspectus. In M. Shibatani (Ed.), Syntax and semantics, vol. 6: The grammar of causative constructions (pp. 1–40). New York: Academic Press.

Shultz, T. R. (1982). Rules of causal attribution. Monographs of the Society for Research in Child Development 47: 1-51.

Sloman, S. (2005). Causal models: How people think about the world and its alternatives. Oxford: Oxford University Press.

Sloman, S. A., Barbey, A., and Hotalling, J. (in press). A causal model theory of the meaning of cause, enable, and prevent. Cognitive Science.

Song, G. (1996). Causation, adicity and lexical aspect. In M. Przezdziecki & L. Whaley (Eds.), ESCOL '95 (pp. 299–307). Ithaca, NY: CLC.

Spelke, E. S., Phillips, A. T., and Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, A. Premack (Eds.), Causal cognition: A multidisciplinary debate (pp. 44-78). Oxford: Oxford University Press.

Spellman, B. A., and Mandel, D. R. (1999). When possibility informs reality: Counterfactual thinking as a cue to causality. Current Directions in Psychological Science 8: 120-123.

Spellman, B. A., Kincannon, A. P., and Stose, S. J. (2005). The relation between counterfactual and causal reasoning. In D. R. Mandel, D. J. Hilton, and P. Catellani (Eds.), The psychology of counterfactual thinking (pp. 28-43). London: Routledge Research.

Talmy, L. (1988). Force dynamics in language and cognition. Cognitive Science 12: 49-100.

Thinès, G., Costall, A., and Butterworth, G. (1991). Michotte's experimental phenomenology of perception. Hillsdale, NJ: Erlbaum.

White, P. A. (1999). Towards a causal realist theory of causal understanding. American Journal of Psychology 112: 605–642.

White P. A. (2006). Theoretical notes: A causal asymmetry. Psychological Review 113: 132 – 147.

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. Cognition 88: 1-48.

Wolff, P. (2007). Representing causation. Journal of Experimental Psychology: General 136: 82-111.

Wolff, P. (2008). Dynamics and the perception of causal events. In T. Shipley and J. Zacks (Eds.), Understanding events (555 – 588). Oxford University Press.

Wolff, P., and Zettergren, M. (2002). A vector model of causal meaning. In W. D. Gray and C. D. Schunn (Eds.), Proceedings of the twenty-fourth annual conference of the Cognitive Science Society (pp. 944-949). Mahwah, NJ: Lawrence Erlbaum.

Woodward, J. (2003). Making things happen. Oxford: Oxford University Press.

Woodward, J. (2006). Sensitive and insensitive causation. Philosophical Review 115: 1-50.

Woodward, J. (2007). Interventionist theories of causation in psychological perspective. In A. Gopnik and L. Schulz (Eds.), Causal learning: Psychology, philosophy, and computation (pp. 19-36). Oxford: Oxford University Press.