Chapter

THE BIAS TOWARD CAUSE AND EFFECT

George Newman*

Yale University, New Haven, CT, USA

THE BIAS TOWARD CAUSE AND EFFECT

There is little doubt that causal reasoning plays a vital role in human cognition. Indeed, it is difficult to see how we could make sense of almost any aspect of the physical or social world without an appreciation of cause and effect. For example, a causal interpretation of the world affords the ability to both make sense of previous events and, in turn, make predictions about future ones – two processes that are clearly valuable to the survival of any organism. As such, causal reasoning has been a major topic of inquiry within philosophy and psychology, encompassing broad questions about the nature of our causal representations, the application and scope of causal reasoning, and the origins of this ability.

Philosophical theories of causality have typically addressed the topic directly — asking metaphysical questions about how (or whether) this property actually exists. Though for centuries philosophers intuitively knew the importance of cause and effect, it was typically assumed that causal relations themselves were readily apparent in the environment and all one had to do was simply observe

* Please address correspondence to: George E. Newman, Assistant Professor of Organizational Behavior, School of Management, Yale University, 135 Prospect St., New Haven, CT 06520, Phone: (203) 432-1046, Email: george.newman@yale.edu

them (Hitchcock, 2003). Hume (1748), however, famously argued that this was, in fact, not the case. By carefully analyzing situations in which causality is commonly assumed, Hume demonstrated that the ‘cause’ itself is nowhere to be found. For example, he argued that relations of cause and effect are not supported even in simple mechanical interactions of two billiard-balls colliding: “Motion in the second Billiard-ball is a quite distinct event from motion in the first; nor is there any thing in the one to suggest the smallest hint of the other” (1748/1977, p. 18).

Hume’s observations drastically changed the nature of the inquiry, such that philosophical questions regarding causal reasoning were no longer questions solely about the world, but rather, could be construed as questions about the mind. The notion that cause and effect can be thought of as a *lens* through which the world is observed has opened the door for a rich and exciting program of psychological investigation, which has flourished over the last century and exploded in the last twenty years.

MAKING THE CASE FOR ‘CAUSE’

One of the primary challenges for psychology has been to demonstrate that the concept of “cause” is a unique concept that is distinct from associative learning. As Waldmann et al. (2006) note, “According to many learning theories, causal predictions are driven by associative relations that have been learned on the basis of observed covariations between events. Similar to Pavlov’s dog who has learned to predict food when it hears a tone (i.e., classical conditioning), or to a rat’s learning that a lever press produces food (i.e., instrumental conditioning), we learn about causal relations” (2006, p. 307).

Though a comprehensive review of all associative accounts is beyond the scope of this paper, in many cases it may actually be quite difficult to distinguish causal interpretation from associative learning—after all, it was Hume’s initial observations that argued for the importance of temporal associations in setting up the ‘illusion of cause.’ Thus, historically, the burden of proof has been to demonstrate that the mind does in fact traffic in the concept of cause, rather than mere associations.

Some of the most convincing evidence for casual reasoning as a distinct sort of mental phenomena comes from the seminal work of Albert Michotte (1963), who demonstrated the primacy of causal interpretations in perception. Using an ingenious apparatus for the time, Michotte presented subjects with simple 2-D displays in which one disc appeared to move until it was adjacent to a second disc, at which point the first disc stopped and the second disc began moving along the same trajectory. Michotte observed that such displays are perceived in terms that transcend their spatiotemporal dynamics. When we see such an event we do not simply see the cessation of one disc's motion followed by the onset of motion in the other. Instead, we see one disc cause the other one to move. Michotte further noticed (in over 100 studies reported in his book) that the perception of causality seemed to be determined by a highly constrained collection of visual cues, and was largely unaffected by higher-level beliefs or intentions. Perhaps most convincingly, Michotte demonstrated that causality is perceived even when the collision itself is physically impossible, for example, when a real 3-D object "collides" with a 2-D light, or when a faster moving object collides with a slower moving object, but "causes" the second object to slow down, rather than to speed up.

Because we perceive causality even in these physically impossible cases, Michotte argued that the perception of cause in his displays could not arise via associations to real-world events. Instead, he concluded that such perceptions must arise from some type of innate causality detector, which responds reliably to actual causal relations, but also approximately to events that resemble real-world causation (for instance, along spatiotemporal dimensions). Indeed, Michotte's speculations about innateness have been supported empirically (e.g., Leslie & Keeble, 1987; Newman, Choi, Wynn & Scholl, 2008; also see Saxe & Carey, 2006). For example, by at least 7 months of age, infants seem to appreciate cause and effect relations via sensitivity to agentive and receptive roles (Leslie, 1986; 1988; Leslie & Keeble, 1987).

This example is but one from a large literature, which has demonstrated the subtle and nuanced nature of causal reasoning, using a wide-range of methods among several different species. A general conclusion that can be drawn from this body of work is that our causal impressions of the world seem to reflect automatic processing that is insulated to some degree from other aspects of cognition. For example, we can see causality in some situations even when we are certain it does not exist (as when a real billiard ball appears to cause a patch of light to move), but we will fail to perceive causality in many other situations when we know that it does exist. Thus, although Hume raised the possibility that 'cause' is merely an illusion, the data suggest that a commitment to a causal interpretation of the world seems to be irresistible and foundational to the structure of the mind.

THE SAME UNDERLYING CONCEPT?

One view of causal reasoning might suggest that the concept of cause and effect is largely the same across different domains and modalities. For example, we may use the term ‘cause’ to describe each of the following phenomena: one ball *causes* the other to move, Jack *caused* Jill to be sad, the bite *caused* his legs to itch. However, it is not immediately clear that the “cause” in each of these cases is in fact, the same. Each situation actually seems to be quite distinct, in terms of the types of features that are involved, the modality, the number of causal links, the abstractness or concreteness of mechanism, etc. Indeed, it may be very difficult to reduce all of the ways in which we understand the notion of cause and effect down to a single definition or a single mathematical model, precisely because the different ways in which perceive cause and effect relationships entail radically different sorts of assumptions and rely upon very different types of knowledge.

Nevertheless, there are a number of theoretical approaches which have attempted to reduce our causal representations to a single, underlying model. Perhaps the approach that has received the most attention in recent years is Bayesian modeling (e.g., Gopnik et al., 2004; Pearl, 2000; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003; Tenenbaum, Kemp, & Shafto, 2007). Bayesian models along with others (e.g., Cheng & Novick, 1990; 1992), essentially start with the same set of assumptions -- causal representations themselves must be built from the bottom-up, with information about the relative probabilities of occurrence serving as the underlying foundation (i.e., covartiation data). The relative weightings of different types of features may vary from theory to theory based on information about causal power (Novick & Cheng, 2004; Rips, in press), but essentially it is the compilation of these statistics that is doing the real cognitive work (also see Shanks & Dickinson, 1987; Wasserman, et al. 1996).

What differentiates the Bayesian approach is that probabilistic information serves as the basis for forming a theory about a particular set of variables and subsequently, the theory is tested and pruned via intervention (or observing interventions by others) on the causal system. In the end, the learner arrives at something like a fully articulated causal theory—for example, knowing all the parts of a CD player and how they work (Rips, in press).

This, of course, is an oversimplification. Moreover, some computer models based on a Bayesian approach have emulated subject data with startling degrees of accuracy (e.g., Griffiths, Kemp, & Tenenbaum, in press). At the same time, there is no shortage of critiques of the Baysian modeling. For example, many Bayesian models rely on an assumption known as the “Markov principle,” which posits that knowledge about a parent cause and its effects renders that unit independent from causes that are further up the causal chain — in other words, to make a prediction about some effect, a learner should only need to focus on the presence or absence of the most immediate cause (Pearl, 2000). However, when tested empirically, people do not seem to obey this principle (Rips, in press). For example, when reasoning about natural kinds, people tend to additionally assume the presence of an unobservable cause, such as an "essence" (Keil 1989; Newman & Keil, 2008; Rheder & Burnett, 2005; also see Burnett & Keil, in prep, for additional critiques of the Bayesian approach).

The goal of the present paper, however, is not to critique Baysian models per se. Rather, I argue that Bayesian models, along with other current theories of causal representation, have perhaps missed the mark in a broader sense. Specifically, these approaches have inadequately addressed two key features of causal reasoning: (1) the role of domain-specific knowledge, and (2) the incompleteness of people's actual causal theories. In turn, I argue that there are in fact many different ways in which we may go about tracking causal structure in the world, and as a result, future research on causal reasoning may be best served by a "taxonomic" approach that attempts to systematically classify each of these different notions of cause, rather than a quantitative approach that tries to reduce all cause and effect relationships to a single formalized structure. The remainder of this paper attempts to outline this argument in greater detail.

CHALLENGES TO A SINGLE NOTION OF CAUSE

Computational models of causal learning typically assume that causal representations are built from the bottom-up. In brief, a leaner (e.g., an adult, or child, or rat) observes the world around them, notices the relative frequency of different events, forms hypotheses about how those events are related, and tests those hypotheses by intervening on the causal system. At the end of the learning process, the learner arrives at a model that incorporates all of these statistics into some kind of coherent and meaningful representation (often depicted as a diagram with nodes and arrows). This model is what allows the learner to make predictions about the future – e.g., whether the Blicket will turn on, what happens when the laser is removed from a CD player, whether or not a Lake Victoria shrimp will gain weight. What is critical here is that the model incorporates all of the relevant features, how they are connected, and some information about their relative rates of occurrence.

However, there is reason to think that these are not the beginning and end-points of the causal reasoning "spectrum"—or at the very least, not the only end-points that exist. First, consider the cognitive developmental literature that has focused on the nature and extent of domain-specific knowledge. Often, these studies are designed to assess whether children at a given age appreciate that one or a few critical features differ drastically between different ontological domains. For example, in a well-known series of studies in which children are told about a raccoon that was cosmetically transformed to appear like a skunk, they report that despite a skunk-like appearance, the animal is still a raccoon. In contrast, when children are told about analogous transformations done to artifacts, such as a coffeepot that is changed to look like a birdfeeder, they report that the object has changed categories (Keil, 1989). This result demonstrates in young children our common-sense conception that biological kinds (but not artifacts) can remain a member of their original category even if their surface appearance changes radically.

There have been a few attempts to integrate (or at least acknowledge) the role of this type of domain-specific knowledge in bottom-up learning. However, when computationally successful, these models have often over-simplified the true nature of people's intuitive theories. For example, they have been characterized by their own authors as "minimalist accounts, intended to capture only those aspects of theories relevant for the basic inductive inferences we study" (Tenenbaum et al., 2007, p. 5). Moreover, there is the larger worry that 'building in' domain-specific knowledge as a set of priors begs the question. It would seem that knowledge about a domain, its members, and relevant features, should also fall under the umbrella of causal reasoning. Put differently, even if domain-specific knowledge can be incorporated into bottom-up learning models, how do we get domain-specific knowledge in the first place?

This concern is even more pressing as numerous studies have confirmed domain-specific knowledge in infants as young as 5 months of age (e.g., Kuhlmeier, Bloom, & Wynn, 2004). Hence, whatever processes govern the formation or maintenance of domain-specific knowledge are in place very early on in life. Rather than trying to fit domain-specific knowledge into the framework of bottom-up learning, an alternative approach may be to simply acknowledge the existence of certain, highly salient relationships that are acquired very early on in life (perhaps with the aid of innate learning mechanisms). For example, the very existence of domain-specific knowledge in infancy suggests other mechanisms by which causal or 'causal-like' knowledge may be acquired.

The second criticism targets the notion that people possess causal models that resemble something like a fully articulated theory. Strong evidence against this idea comes from work on the Illusion of Explanatory Depth (IOED) (e.g., Lawson, 2006; Mills & Keil, 2004; Rozenblit & Keil, 2002). In short, this research finds that people tend to overestimate their understanding of complex causal phenomena. For example, if people rate how well they understand the workings of a helicopter, are then asked to provide a detailed description of how a helicopter works, and then re-rate their knowledge in light of their explanation, their ratings drop dramatically – evidently because have been confronted with the true incompleteness of their theory (Rosenblit & Keil, 2002). Evidence of this illusion of depth has been documented in several different knowledge domains and age ranges (Mills & Keil, 2004) and is distinct from more general biases towards overconfidence (Rozenblit & Keil, 2002).

The IOED may occur for a number of reasons, such as people's tendency to underestimate the true complexity of various causal phenomena. However, more recent research suggests that the roots of the phenomenon may actually run far deeper: illusions of knowledge depth seem to derive from the fact that people's actual causal models are, for the most part, constructed 'on-the-fly' to accommodate the task at hand. For example, when searching to provide an explanation for how a common device works, people will often construct models that incorporate features that have been erroneously added to the stimulus set (Keil, in prep). Thus, the idea that people possess something like a stable, coherent causal representation appears to be somewhat of a misnomer. Instead, people's actual causal models seem to be quite sparse, consisting of knowledge of only a few key features and a vague notion of cause.

Does the existence of an illusion of explanatory depth in fact undermine notions of a fully articulated theory? One could argue "no," that such effects merely reflect a lack of knowledge about certain processes in certain domains and perhaps overconfidence about how much one knows. However, the data suggest that in terms of folk-knowledge, it is not clear that such sparse representations yield radically different inferences from more complete causal theories, or that these sparse representations are completely lacking in structure. For example, Rozenblit & Keil (2002) note the similarities between an illusion of depth and traditional interpretations of naïve essentialism, where it is typically assumed that people conceptualize essences more as causal "placeholders," rather than as actual physical entities (Gelman, 2003; Medin & Ortony, 1989). Thus, the fact that people seem to possess relatively sparse ad hoc causal models, rather than fully articulated theories, does not seem to undermine the inductive potential of people's intuitive reasoning. It does, however, challenge the notion that people arrive at fully-articulated theories that are built from the bottom up.

A DIFFERENT APPROACH

In his paper titled, "Of Humean Bondage," Christopher Hitchcock (2003) highlights philosophers' use of "attachment" metaphors to convey different notions of cause. Perhaps the most salient, is the metaphor of 'cause as glue' – i.e., as a physical substance that binds two concepts together. Hitchcock goes on to critique any literal use of this metaphor – e.g., the notion that causes are in fact bound to their effects, such that causal relationships can be tested via counterfactual reasoning. Moreover, through several clever thought experiments, Hitchcock essentially dispels the notion that a formal and complete definition of 'cause' can be constructed. For example, he demonstrates how in many cases, it may actually be quite difficult to definitively say that one thing has caused another.

Hitchcock's sentiment is echoed by other critiques in the psychological literature, which have argued that existing psychological theories are unable to account for nuances evident in human causal reasoning (e.g., Rips, in press). For example, as I argued earlier, domain-specific knowledge or effects such as the IOED present a serious challenge for traditional bottom-up conceptions of causal learning. Similarly, numerous studies suggest a form of causal reasoning that is quite early emerging, highly sensitive to domain-specific constraints and capable of profound developmental change, even in the absence of knowledge about actual causal mechanism (Newman & Keil, 2008).

In other words, the assumption that the concept of cause and effect is one particular kind of thing with a discernable internal structure (i.e., a certain kind of mental glue), has not been entirely successful. In many places, we are left wondering how this same notion of cause could do so much: for example, in imposing highly specific constraints on perception, drawing on broad divisions between ontological domains, and undergoing radical developmental shifts at a level far above concrete mechanism.

Alternatively, I suggest that it may be more beneficial to approach the topic of causal reasoning not by searching for the underlying composition of "causal glue," but rather by seeking to identify and categorize distinct types of causal thinking. In analyzing the utility of folk-physics, Elga (2007) essentially concludes that our folk-intuitions, which emphasize some forces like gravity and ignore others like nuclear forces, are metaphysically sufficient because the physical forces that operate in our immediate environment cohere as an isolated cluster that is immune from more distant forces. Thus, even though what happens 20 million light years away should in theory impact what happens on earth (according to lawful determinism), in practice, the forces are simply to distant and weak to matter. To extend this metaphor somewhat further, one could argue that there may be many different domains in which causal relations can be similarly thought of as coherent units that are functionally isolated. For example, consider our folk-psychological intuitions. Though it is of course true that thinking is the product of physical matter, which is subject to the same physical laws as all other matter, practically speaking, when we try to interpret or predict the behavior of others, it is incredibly important to consider only those factors that tend to cohere as part of folk-psychology (intentionality, goal-directedness, etc.) and ignore those that do not (e.g., the neurobiology of the brain, or the behavior of individual neurons).

In other words, one could think of causal relations in the world as actually many different isolated clusters that each consist of their own set of features and causal assumptions. It is important to note that this conception of causal knowledge has strong parallels to the way in which "intuitive theories" are typically discussed (e.g., Keil, 2007; Murphy & Medin, 1985). It may be that there are several different kinds of these stable causal clusters in the world and we are sensitive to them, but perhaps in different ways depending on the kind of cluster. Each of these clusters may be distinct either for reasons that are specific to the functioning of the mind (i.e., we tend to process *this* collection of stuff in the same way), reasons that are specific to the structure of the world (i.e., the collection has meaningful coherence in reality), or in a very "Gibsonian" sounding way, reasons that reflect the elegant interaction between the two.

Moreover, the way in which we represent these different clusters could of course be very different. Some clusters may be best understood more in associative terms, while some may be represented more theoretically; some representations may be relatively sparse, while others may be incredibly rich. Thus, while the overarching term 'cause' may provide a rough heuristic for identifying agentive and receptive roles in the environment, the deeper nature of cause may vary widely from one causal cluster to the next. Under this view, studying causal reasoning as the reflection of a single process or single type of concept seems much less productive then studying it as *many things*, each with their own type of internal structure.

In turn, the role of cognitive science and the yield from past and current empirical work may instead be to map out the taxonomy, position and the relation of these different ways of tracking causal structure in the world. Here I think we have made considerable progress, as different literatures within the field tend to cohere over similarities in subject population, methodology, or topic. At the same time, it will be important to consider each of these factors and the exact conditions under which different empirical results have been obtained – for instance, whether the subjects are adults or infants, whether the learning phase is extensive or minimal, or whether the stimulus is a Blicket or a Brine shrimp. It may be that careful attention to even these superficial similarities will reveal a great deal if one is looking for many different types of causal reasoning, rather than just one.

In time, analogous to the historical progression from astronomy to physics, deeper commonalities may become more apparent as a more detailed hierarchy is constructed and functional similarities give way to computational or neurological ones. It may well be that ultimately all causal representations can in fact be linked under a single, unified explanation. I argue, however, that this remains to be seen only with a different type of theoretical approach. In the meantime, the study of causal reasoning is alive and well and the wealth of data generated over this topic provides a substantive basis for charting the landscape of causal concepts. At the same time, it may be best (at least temporarily) to dispel the idea that one single theory will encapsulate all representations of cause and effect. Instead, it may be more theoretically productive to begin with the assumption that there are in fact many distinct notions of cause (each with their own type of representation), and then map out the functional similarities and differences between them.

References

Blaisdell, A. P., Sawa, K., Leising, K. J., & Waldmann, M. R. (2006). Causal reasoning in rats. *Science*, *311*, 1020-1022.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, *58*, 545-567.

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

Choi, H., & Scholl, B. J. (2006). Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception. *Perception*, *35*, 385-399.

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology*, *29*, 421-433.

Cohen, L.B. (1998). An information-processing approach to infant perception and cognition. In F. Simion & G. Butterworth (Eds.), *The Development of Sensory, Motor, and Cognitive Capacities in Early Infancy* (pp. 277-300). East Sussex: Psychology Press.

Cohen, L. B., Amsel, G., Redford, M. A., & Casasola, M. (1998). The development of infant causal perception. In A. Slater (Ed.), *Perceptual development: Visual, auditory, and speech perception in infancy* (pp. 167-209). East Sussex, UK: Psychology Press Ltd.

Cohen, L.B., & Amsel, G. (1998). Precursors to infants' perception of causality. *Infant Behavior and Development*, *21*, 713-731.

Elga, A. (2007). Isolation and folk physics. In Huw Price & Richard Corry (Eds.) *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited* (pp. 106-119). Oxford University Press.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford: Oxford University Press.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3-32.

Griffiths, T. L., Kemp, C., and Tenenbaum, J. B. (in press). Bayesian models of cognition. To appear in Ron Sun (ed.), *Cambridge Handbook of Computational Cognitive Modeling*. Cambridge University Press.

Hitchcock, C. (2003). Of Humean Bondage. *The British Journal for the Philosophy of Science, 54*, 1 -25.

Hume, D. (1748/1977). *An enquiry concerning human understanding*. Indianapolis: Hackett, 1977.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge: MIT Press.

Keil, F. C. (2007). Biology and Beyond: Domain Specificity in a Broader Developmental Context. *Human Development*, *50*, 31-38.

Kuhlmeier, V.; Bloom, P.; & Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition*, *94*, 95-103.

Lawson, R. (2006). The science of cycology: Failures to understand how everyday objects work. *Memory & Cognition, 34*, 1667-1675.

Leslie, A. M. (1986). Getting development off the ground: Modularity and the infant's perception of causality. In P. van Geert, (Ed.), Theory building in development (pp. 405 - 437). Amsterdam: Elsevier, North-Holland.

Leslie, A. M. (1988). The necessity of illusion: Perception and thought in infancy. In L. Weiskrantz (Ed.), Thought without language (pp. 185-210). Oxford: Clarendon Press/Oxford University Press.

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

Levin, D. T. & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures, *Psychonomic Bulletin and Review, 4*, 501-506.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical learning* (pp. 179–195). New York: Cambridge University Press.

Michotte, A. (1946/1963). *La perception de la causalite*. Louvain: Institut Superior de Philosophie, 1946. English translation of updated edition by T. Miles & E. Miles, *The perception of causality*, Basic Books, 1963.

Mills, C. and Keil, F .C. (2004). Knowing the limits of one's understanding: The development of an awareness of an illusion of explanatory depth. *Journal of Experimental Child Psychology, 87*, 1-32.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*, 289–316.

Newman, G. E., Choi, H., Wynn, K., & Scholl, B. (2008). The origins of causal perception: Evidence from postdictive processing in infancy. *Cognitive Psychology.*

Newman, G. E., & Keil, F. C. (2008) Where's the essence? Developmental shifts in children's beliefs about internal features. *Child Development.*

Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal power. *Psychological Review*, *111*, 455-485.

Pearl, J. (2000). *Causality*. Cambridge, UK: Cambridge University Press.

Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology*, *50*, 264-314.

Rensink, R. A.; O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: the need for attention to perceive changes in scenes, *Psychological Science, 8*, 368-373.

Rips, L. J. (in press). Causal thinking. In J. E. Adler & L. J. Rips (Eds*.), Reasoning: Studies of Human Inference and Its Foundation*. Cambridge, UK: Cambridge University Press.

Rozenblit, L. R. & Keil, F. C. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science, 26,* 521-562.

Saxe, R., & Carey, S. (2006). The perception of causality in infancy. *Acta Psychologica*, *123*, 144-165.

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. *Psychology of Learning and Motivation*, *21*, 229-261.

Simons, D. J. & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction, *Psychonomic Bulletin and Review, 5*, 644-649

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observation and interventions. *Cognitive Science*, *27*, 453-489.

Tenenbaum, J. B., Kemp, C., & Shafto, P. (2007). Theory-based Bayesian models of inductive reasoning. In A. Feeney & E. Heit (Eds*.), Inductive reasoning*. Cambridge University Press.

Waldmann, M. R., Hagmayer, Y, & Blaisdell, A. P. (2006). Beyond the information given: Causal models in learning and reasoning. *Current Directions in Psychological Science*, *15*, 307-311.

Waldmann, M. R., & Hagmayer, Y. (2006). Categories and causality: The neglected direction. *Cognitive Psychology, 53,* 27-58.

Wasserman, E. A., Kao, S. F., Van Hamme, L. J., Katagiri, M., & Young, M. E. (1996). Causation and association. *Psychology of Learning and Motivation*, *34*, 207-264.