PART III
Psychology

·7·
Causal thinking

David Lagnado

Abstract

How do people acquire and use causal knowledge? This chapter argues that causal learning and reasoning are intertwined, and recruit similar representations and inferential procedures. In contrast to covariation-based approaches of learning, I maintain that people use multiple sources of evidence to discover causal relations, and that the causal representation itself is separate from these informational sources. The key roles of prior knowledge and interventions in learning are also discussed. Finally, I speculate about the role of mental simulation in causal inference. Drawing on parallels with work in the psychology of mechanical reasoning, the notion of a causal mental model is proposed as a viable alternative to reasoning systems based in logic or probability theory alone. The central idea is that when people reason about causal systems they utilize mental models that represent objects, events or states of affairs, and reasoning and inference is carried out by mental simulation of these models.

You arrive at your holiday apartment, welcomed by the local cat and a chorus of crickets outside the window. During the night your sleep is interrupted by intermittent high-pitched squeals. At first you assume it is the cat, but on careful listening the noises sound mechanical rather than animate. You get up and walk around the flat, and notice that the light on the smoke detector is flashing red – suggesting that the battery is running down. You recall a similar problem with the smoke alarm in your own house, and an equally annoying high-pitched squeal as the battery died out. You remove the battery from the fire alarm, and the squeals stop.

Next morning, as you make breakfast, the squeals seem to return. But the smoke detector is lying dismantled on the table. Perhaps the capacitor is still discharging, emitting the occasional squeal? Or a cricket has started mimicking the sound of the dying smoke detector? But then you notice that whenever you turn on the kitchen tap, there is a high-pitched noise that sounds very similar to last night's squeals. You turn on the tap at random moments throughout the day, and it is nearly always followed by a squeal. Problem solved! But maybe the smoke detector, like the local cat, was falsely accused. Perhaps it was the dodgy plumbing all along? Just as you start to reinsert the battery to test out this idea you remember that you are on holiday.

978–0–19–957413–1 07-Mckay-Illari-c07-drv Mckay (Typeset by SPi, Chennai) 130 of 928 October 30, 2010 12:47
OUP UNCORRECTED PROOF – FIRST PROOF, 30/10/2010, SPi

This is an everyday example, but it illustrates several of the key aspects of causal thinking. In this chapter I will use it as a running example to identify some shortcomings with current theorizing in the psychology of causal inference. I will also suggest ways in which our understanding of causal cognition might be improved. In particular, I will argue that causal learning and reasoning are intertwined, that there are multiple sources of evidence for causal relations, and I will speculate about the role of mental simulation in causal inference.

7.1 Interplay of learning and reasoning

At a general level the smoke detector example highlights the interplay between causal learning (typically conceived as the induction of causal relations from patterns of observations) and causal reasoning (drawing inferences on the basis of assumed causal relations) – and the artificiality of separating and studying these activities in isolation, as is common in the psychological literature. Thus, as you try to work out the cause (or causes) of the high-pitched squeals, you engage in a variety of interleaved inferential activities, including hypothesis generation and testing, hypothetical and counterfactual reasoning. For example, your observation that the smoke detector light is flashing red leads you to hypothesize that the battery is running low, and that a low battery is the cause of the noises. You reason (hypothetically) that if the low battery is the cause, then removing the battery altogether should stop the squeals. You confirm this hypothesis by removing the battery and noting that the squeals stop.

Even in this simple learning episode, which is only the start our scenario, you have deftly switched between various forms of inference. You have inferred a putative causal relation from observations, but also engaged in hypothetical reasoning and hypothesis-testing. In mainstream cognitive psychology, however, causal learning and causal reasoning are usually treated as separate areas of research, with different theories and empirical paradigms.

Research in causal learning focuses on the induction of causal relations from data, with little concern for the other reasoning activities that might accompany this induction. A typical experiment presents people with covariation information about potential causes and effects (either in summary format, or presented sequentially) and asks them to assess the strength of the putative causal relation. Applied to our example, this would correspond to showing people a sequence of cases in which the smoke detector battery is either low or high, and the detector either does or does not make a squeal. The central question of interest to experimenters and theorists is how people arrive at a causal estimate from the patterns of covariation, and there is substantial debate about this (Cheng, 1997; Griffiths & Tenenbaum, 2005; Shanks, 2004; Vallee-Tourangeau *et al.* 1998). However, very little is said about what, if any,

reasoning occurs in such experiments, or how this reasoning takes place. Indeed theorists of an associative persuasion maintain that people are simply acquiring associations between mental representations of the presented variables (e.g. mentally associating 'low battery' with 'squeals')[1]. But this is only part of the story. There is compelling empirical evidence that in causal learning contexts people are not merely associating the two variables, but are hypothesizing that one *causes* the other (Waldmann, 1996; Waldmann and Holyoak, 1992; see Lagnado *et al.* 2007 for a review), an inference that engenders a range of further inferences (e.g. that if you were to replace the low battery with a new one, then the squeals will stop; and if the battery had been high, then there would have been no squeals). Of course these inferences are fallible. You might have the wrong causal model. But your conjecture that one variable causes another carries with it these additional inferences, in a way that the postulation of a mere association does not. The claim that two variables are associated by itself tells us nothing about what would happen to one variable if we were to change the other.

More generally, an analysis of causal learning that focuses only on the associations that people can learn between variables does not account for a variety of other inferential activities that people engage in, or the wealth of information (beyond covariation data) that they might use to support these inferences.[2]

On the other hand, research in causal reasoning tends to focus on how people make conditional or counterfactual inferences on the basis of presupposed causal relations, with little regard for how these causal models are acquired or generated. For example, one of the dominant accounts of reasoning, mental model theory, argues that causal reasoning involves the construction of possible states of the world and the search for counterexamples (Goldvarg and Johnson-Laird 2001). In particular, the meaning of 'A causes B' is cashed out in terms of the possibilities consistent with 'A materially implies B' and the constraint that A temporally precedes B. However, nothing is said about how people acquire the background causal knowledge that allows them to generate appropriate possibilities, and make appropriate connections between these possibilities (e.g. distinguish between possible states that are causal connected rather than merely correlated).

Other investigations into causal reasoning (e.g. explanation-based reasoning, Hastie & Pennington 2000; causal heuristics, Kahneman & Tversky 1982) also fail to account for how people construct and revise causal models. This would not be a problem if learning and reasoning were separate activities that engaged entirely distinct processes; but these inferential activities are

[1] There are several varieties of associative theory, but this appears to be a shared assumption across most variations (Dickinson, 1980; Hall, 2002; Shanks, 1995)

[2] This is not to undermine the important role that associative learning can play in cognition, but to emphasize that causal thinking will often go beyond the acquisition of associations.

978–0–19–957413–1 07-Mckay-Illari-c07-drv Mckay (Typeset by SPi, Chennai) 132 of 928 October 30, 2010 12:47
OUP UNCORRECTED PROOF – FIRST PROOF, 30/10/2010, SPi

best studied together, as part of a general system of causal inference (see Lagnado 2009). Postulating common representations and mechanisms also explains why both learning and reasoning are subject to similar constraints (e.g. attentional or working memory limitations).

From a formal perspective, causal bayesian networks (Spirtes, Glymour and Scheines 1993; Pearl, 2000) provide a unified framework for representation, learning and inference (for criticisms of some of the assumptions underlying this framework see Cartwright, 2007, Williamson, 2005). Numerous psychologists have adopted this framework as a model of human inference (Gopnik and Schultz, 2007; Lagnado *et al*. 2007; Sloman and Lagnado, 2005; Sloman *et al*. 2009; Tenenbaum *et al*. 2007). The framework is a great advance insofar as it provides a normative benchmark against which to appraise human inference, and a guide for the construction of descriptive models. But the formalism alone does not mandate any specific psychological account; and, indeed, a psychological theory need not be tied too tightly to the normative theory.

7.2 Multiple sources of evidence for causal beliefs

Psychological research into causal learning has been dominated by the question of how people induce causal relations from patterns of covariation. However, covariation-based learning is only part of the picture, and exclusive focus on this question threatens to distort our understanding of causal cognition.

As well as engaging in several kinds of inference (not just induction), people use various sources of information to infer causal relations (Einhorn and Hogarth, 1986; Lagnado *et al*. 2007; Waldmann *et al*. 2006). These 'cues to causality' include information about temporal order, interventions, spatiotemporal contiguity, similarity, analogy and prior knowledge. The smoke detector scenario in fact illustrates most of these possibilities. For example, the temporal proximity between squeals and tap turns was an important clue to identifying the hot water system as a likely cause (if the temporal interval had been much greater, or more variable, you would have been less likely to associate the two). The repeated interventions on the hot tap, at a random selection of times throughout the morning, helped to rule out possible confounding causes, and establish the hot tap as a cause of the squeals. The analogy of the current situation to a previous encounter with the noises admitted from a smoke detector helped guide the hypothesis formulation and testing. The similarity in sound of the squeals meant that a single cause was sought (e.g. smoke detector or hot water system). The role of prior information was also ubiquitous (see next section).

This is not to deny the role of covariation information (which also plays its part in our scenario), but to emphasize that it is just one cue among many. Indeed covariation information by itself is seldom sufficient for inferring a

unique causal model. For example, consider a variation on our story about the mysterious squealing where we ignore the smoke detector and faulty hot water system, and focus on the local cat. Suppose that the only available evidence is your observation of a strong correlation between the appearances of the cat and the sound of the squealing noises. In the absence of any other information (e.g. about temporal order; spatiotemporal contiguity etc.) this evidence does not distinguish between a causal model in which the cat causes the squeals, or a model in which the squeals cause the cat's presence (perhaps it is tracking a squealing mouse or large insect, or even a mate), or a common cause model in which both the cat's presence and the squeals are due to a third factor.

One way to distinguish between these models is to intervene by removing the cat from the premises, and ensuring it does not return. Do the intermittent squeals persist? If so, then the cat is ruled out as a cause. If not, then it is ruled in. Another route is to acquire additional information about the potential cat $\rightarrow$ squeal relation, perhaps in terms of spatial or temporal information about the cat and the squeals. The critical moral is that the mere observation of a correlation does not provide evidence for a unique causal relation (for more details see Lagnado *et al*. 2007; Sloman, 2005).

Recent psychological studies have shown that people are indeed able to use interventions to learn causal models; and, moreover, to learn models that cannot be learned from covariational information alone (Gopnik *et al*. 2004; Lagnado and Sloman, 2002, 2004, 2006; Meder *et al*. 2008; Steyvers *et al*. 2003; Waldmann & Hagmayer, 005). Not only do people learn better when they can intervene on a system, but they also make appropriate inferences about the effects of their interventions (Sloman and Lagnado, 2005). These experiments, as well as numerous others conducted by Tenenbaum and Griffiths and colleagues (e.g. Tenenbaum and Griffiths, 2003; Tenenbaum *et al*. 2007), convincingly demonstrate that people do not solely rely on covariational information to induce causal structure. Instead, they make use of a range of sources of information, including a central role for the evidence gathered from interventions.

In many real-world contexts people are provided with a rich variety of information about the causal systems they interact with. The control of objects, tools and simple devices (e.g. pens, scissors, can-openers) are readily learned through a combination of interventions, sensorimotor feedback, and spatiotemporal information. To explore this kind of learning we introduced a novel experimental paradigm in which subjects manipulated on-screen sliders in a real-time learning environment. Their task was to discover the causal connections between these sliders by freely changing the settings of one slider and observing the resultant changes in the other sliders. Subjects excelled at this task, rapidly learning complex causal structures with a minimum of exploration (Lagnado *et al*. 2007; Lagnado & Loventoft-Jessen, in prep.).

This set-up revealed two important points about people's capacity for causal learning. First, it only took a few manipulations of a slider for subjects to leap to a causal conclusion about the link between one slider and another. The causal relation 'popped-out' due to the confluence of various factors: the spatiotemporal similarities in the motions of the sliders that were causally connected; the sensitivity of control that one slider exerted on the other; the opportunity for subjects to intervene when they chose (thus ruling out confounding variables).[3] Second, once subjects had explored the system for several minutes, they were able to construct mental models of the causal connections between sliders, and imagine the effects of interventions that they had not yet taken.

This was shown in a follow-up experiment (Lagnado & Loventoft-Jessen, in prep.), in which subjects made use of 'double interventions' in order to disambiguate between models. For example, if subjects are only able to move one slider at a time, it is impossible to distinguish between a model with a chain ($A \rightarrow B \rightarrow C$) and a similar model with an additional link from A to C. In both cases, changes in A lead to changes in B and C, and changes in B lead to changes in C alone. One way to distinguish these models is to disable B, and then see whether changes in A still lead to changes in C. If they do not, then the true model is the chain. In the experiment, subjects engaged in an initial learning phase where they were restricted to moving one slider at a time. At the end of this phase they were asked which causal model (or models) best explained the observations they had made. In a second test they were asked to choose one disabling intervention (in combination with a single slider move) that would allow them to distinguish between models. Many subjects were able to select the correct disabling intervention, showing that they could mentally represent possible causal models, and imagine the effects of interventions on this model (in particular, what would happen if they disabled one slider, and then moved another).

This experiment supports several of the claims made in this chapter. It shows that people can make use of various sources of information, including interventions and spatiotemporal similarities, to learn causal models. It shows that people can engage in hypothetical reasoning in order to disambiguate complex causal structures, thus confirming the interplay between learning and reasoning. Finally, it anticipates the discussion of mental simulation and causal reasoning presented in later sections.

A central claim in this chapter is that it is mistaken to focus on just one source of information, to the exclusion of other sources. A related mistake would be to conflate one source of evidence for causality (e.g. covariation)

[3] These findings have parallels with Michottean paradigms (Michotte, 1954; for recent discussion see Wagemans *et al.* 2006). However, learning in our experiments was not dependent on spatial contiguity, and the causal mechanisms linking the sliders were invisible.

with the conception of causality that people actually have or ought to have.[4] It seems clear from the psychological literature that people's lay conception of causality is rich and multi-faceted, and not reducible to a purely probabilistic notion.

Given that people use multiple sources of information to infer causal beliefs, the question arises as to how this information is combined. In some contexts this will be relatively trivial, because the different cues will converge on the same causal conclusion. This often occurs when agents act in the natural environment – the information given by interventions tend to be nicely correlated with spatial and temporal information – I swat a fly, and the fly dies at a nearby time and place. However, the correlations between cues are sometimes broken – turning the hot water tap causes a squeal to emanate from pipes above my head. Here I use my hazy knowledge of the hot water system to explain this discrepancy. More problematic are cases where two separate cues point in different directions, for example when the temporal ordering suggests one causal model, but the covariation information suggests a different model (see Lagnado & Sloman 2006, for experiments that explore this kind of situation in the context of the appearance and transmission of computer viruses).

The open question is how people combine these different cues, especially when they suggest different causal conclusions. One general approach is to estimate the reliability of each cue, and combine them relative to this weight. However, this might not reflect what people actually do – certain cues in certain contexts might trump other cues altogether (this is what was suggested in Lagnado & Sloman, 2006, with temporal order trumping covariation information). Another possibility is that just as people tend to entertain or test only one causal hypothesis at a time, they also use cues in a sequential fashion. Lagnado and Sloman suggested that people set up an initial hypothesis or model on the basis of time order alone, and then used covariation information to test this model. This was supported by a later experiment that elicited people's causal models regularly throughout the course of the experiment (but more research is needed).

7.2.1 Prior causal knowledge

In most situations causal learning takes place against the backdrop of extensive prior causal knowledge. This knowledge can be very general – that cats sometimes screech, that water systems malfunction, that batteries run low, or more specific – that smoke detectors are battery operated, that the red light on

[4] It could be argued that such a mistake is sometimes made in philosophical circles – especially when theorists attempt to define causation purely in probabilistic terms (Suppes, 1970). Indeed this mistake is perhaps perpetuated in more recent theories of causality based on causal Bayesian networks. Sources of evidence for causal models, whether from observational or interventional probabilities, should not be taken as definitional of causality.

a smoke detector indicates a low battery, etc. This knowledge includes spatial and temporal information – e.g. concerning the relation between location and sound, and mechanical information of all sorts – e.g. the usual functioning of a smoke detector etc. Moreover, people do not require detailed (or correct) knowledge about causal systems in order to use this knowledge to acquire new beliefs. Simple beliefs will often suffice to figure out a novel causal relation. For example, one can infer that the low battery is causing the squeals without detailed knowledge about the inner workings of batteries or smoke detectors (although it helps to know that designers might have constructed the detector so that it warns users when it is about to fail).

A clear illustration of the role of prior knowledge is provided by cases of one-trial learning, where people learn (or assume that they learn) a causal relation after exposure to just one exemplar. For example, in our tale of the smoke detector, it took just one test (in which the battery was removed and the squeals stopped) to establish the hypothesis that a low battery was causing the squeals. It might be argued that this test actually involved a couple of observations – e.g. low battery and squeal, no battery and no squeals. But the point is that the number of observations were definitely too low for standard covariation-based learning algorithms to do their work. Most learning algorithms, including those developed within the CBN framework (e.g. Spirtes *et al*. 1993), require repeated observations before a relation can be learned (in the same way that statistical analyses require datasets larger than one). In cases of rapid learning, prior causal knowledge is used in combination with a simple piece of inferential reasoning.

An associative theorist might respond that one-trial learning can be captured in a contingency-based learning rule, so long as the learning rate parameter is high. In other words, a single observation that the squeals stop when the battery is removed provides enough covariation information to support the causal conclusion that the low battery was the cause of the noises. But this move seems unprincipled, in the sense that one would not want to licence single-trial learning in all contexts. Whether or not one makes the leap to a causal conclusion from just one exemplar (or a few) depends heavily on what other prior background knowledge one has. This inductive leap should only be taken when the background is rich and sufficient to ground the inference (e.g. given basic knowledge about how batteries work; how smoke detectors might be designed, etc.). A single-case co-occurrence in a context where there is little prior knowledge to support the inference, or even knowledge that goes against it, is less likely to lead to rapid learning.

Thus any attempt to address one-trial learning by adjusting the learning parameter in an associative mechanism effectively concedes that additional background information is being used in these cases (and is reflected in the adjustment of this parameter). The crucial point is that it is the background knowledge that is modulating the inferences drawn in one-trial cases, not

the covariation information. This background knowledge supports additional reasoning about the situation — and this explains our ability to learn causal relations from impoverished data (see also Tenenbaum and Griffiths 2003).

There are numerous routes by which people attain prior knowledge – they might have been taught it, read about it, or possibly acquired it firsthand through there own experiences with the causal system in question. The important point is that it is rare for people do be confronted with a causal inference problem for which they have no relevant prior knowledge. Even infants seem to enter the world with certain prior assumptions that help them acquire more specific causal knowledge (Schlottmann, 2001; Scholl, 2005). Despite its ubiquity, the interaction of prior causal knowledge with novel problem situations, and the ability to construct new causal models from prior assumptions, has not been systematically investigated in mainstream cognitive psychology (but see Ahn and Kalish, 2000, Tenenbaum and Griffiths, 2003, Waldmann, 1996).

7.2.2 Prior assumptions behind interventions

One of the key aspects of causal thinking is that it serves as a guide to action. If done right, it allows us to predict and anticipate the effects of our actions, including those that we have never taken before. Pearl (2000) summarizes this neatly with his claim that causal models are 'oracles for interventions'. The flipside of this is that causal models can often be learned through carrying out appropriate interventions. For instance, when I conjecture that the low battery is causing the squeals, I construct a simple causal model: low battery $\rightarrow$ squeal. I then reason that according to this model, if I intervene and replace the old battery with a new one, then the squeals will stop. I can then test this prediction by actually replacing the battery, and observing whether or not the squeals stop. Once I have established the correctness of this causal model, I can use it to make predictions on other occasions. Of course I must be aware that the context might change in ways that make this model inappropriate. There is no guarantee that what works on one occasion will work in the future, or generalize to other slightly different circumstances. I will make assumptions that may have better or worse justifications. Thus, I can safely assume that the same smoke detector will work similarly tomorrow (although I can't be sure – perhaps when I replace the battery another component will break), and also assume that the smoke detector next door operates in the same way. But I will be on dangerous ground if I assume that a very different device (e.g. a battery-operated baby doll) will stop squealing once I replace the battery.

This shows that our causal reasoning depends on assumptions, many of them tacit, about the robustness of the causal models we can learn. Indeed a crucial element in our ability to think causally is our ability to gauge when we can generalize and transpose our models to novel environments

(cf. Cartwright, 2007; Steele, 2007). This seems to be an unexplored area in cognitive psychology.

7.2.3 Causal Bayesian networks over- and under-estimate human reasoners

This concludes our brief survey of the multiple sources of evidence for causal beliefs (for more details see Einhorn & Hogarth 1986; Lagnado *et al.* 2007). One significant point to emerge from this concerns the applicability of the Causal Bayesian Networks (CBN) framework as a model of human inference. A strong advantage for this framework is that it formalizes the distinction between interventional and observational (correlational) learning, and suggests various algorithms for learning causal models under either regime (given certain crucial assumptions). However, there are reasons to question its wholesale application to everyday human causal inference.

In particular, it appears that the CBN framework both over- and under-estimates the capabilities of human reasoners. It seems to over-estimate people's abilities to perform large-scale computations over large bodies of hypotheses and data. People have limited-capacity working memory, and this serves as a bottleneck for complex computations with many variables and relations (Cowan 2001; Halford *et al.* 2007; Miller, 1956). It is likely that human reasoners adopt strategies to overcome these limitations, such as performing local computations (Fernbach & Sloman, 2009) and chunking information into hierarchically structured representations (Bower, 1970; Lagnado & Harvey, 2008).

On other hand, current attempts to apply CBN to human inference also seem to underestimate human capabilities. As noted above, there is a wealth of information about causality aside from statistical covariation, including spatiotemporal information, similarity, temporal order etc. People are able to incorporate this information in their search for causal structure, but this is not yet captured in standard causal Bayesian models. This is not to deny the relevance of the CBN framework, or the considerable benefits it brings to the study of causal cognition. But it is a starting point for formulating better psychological theories, not an endpoint.

7.3 Mental models and simulations

So far we have talked about causal inference at a relatively abstract level, without delving into the mechanics of how people actually carry out these inferences. This level of description is appropriate when comparing people's inferences against a normative model of inference, such as that provided by logic, probability theory or causal Bayesian networks. But it tells us little about the psychological processes that underpin these inferences. For example, if someone's inferences correspond to those prescribed by the normative model,

there remains the question of how these inferences were actually carried out. There will usually be a variety of possible psychological processes that could have reached the normatively correct conclusions.[5]

It is instructive here to compare with the case of deductive reasoning. Some theorists argue that when people make deductive inferences (e.g. from premises 'If X, then Y' and 'X', infer conclusion 'Y') they apply formal inference schema to syntactically structured mental representations (Inhelder & Piaget 1958; Braine and O'Brien, 1998; Rips, 1983). This 'mental logic' theory is controversial, especially in light of the well-documented failures in people's deductive reasoning, and its sensitivity to both content and context (Wason, 1983; Evans, 2002). One alternative to this position is mental model theory (Johnson-Laird, 1983; 2006). On this theory people evaluate deductive inferences by envisaging and combining possible states of affairs. I will not go into the details of this debate. What is important for current purposes is that we should not simply assume that when people reason causally they use a causal logic that operates on sentential mental representations. But what are the alternatives?

Mental model theory provides an alternative perspective here. As noted above, the theory claims that causal reasoning involves the envisioning and combining of possible states (Goldvarg and Johnson-Laird, 2001). There are various reasons why the theory by itself does not seem satisfactory. Prominent amongst these are the lack of constraints on the possible states implied by causal relations (material implication is too inclusive a relation), the failure to account for people's causal judgments (Sloman, Barbey & Hotaling, 2009) and the difficulty the model has in distinguishing inferences based on observations from those based on interventions (for details see Glymour, 2007; Sloman and Lagnado, 2005).

Despite these shortcomings, mental model theory does contain the seeds of a plausible account. To articulate this, it helps to return to the classic work on mental models by Kenneth Craik:

> If the organism carries a 'small-scale model' of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and the future, and in every way react in a much fuller, safer, and more competent manner to the emergencies which face it.
>
> (Craik 1952, p. 61).

Craik's suggestion is that people anticipate the effects of their actions by simulating these actions on a mental model of the external world. The key idea is that manipulations of the mental model parallel the correspondent

[5] This does not mean that conformity to the normative model tells us nothing about the nature of the psychological processes. For instance, successful causal inference presumably requires the capability to represent networks of directed relations between variables.

manipulations of the world. In other words, causal inferences are carried out by mental simulations that somehow encapsulate the causal processes in the world. This proposal is innovative, but raises a host of questions, especially with regard to what constitutes a mental model, and how these models are simulated. For example, what aspects are represented, how these are represented, and how a simulation is actually run so that it respects real world processes. In the following I will pursue one possible extension of Craik's original ideas, one most relevant for causal inference.

7.3.1 Mechanical reasoning

Mechanical reasoning furnishes some clear examples of mental simulation (Hegarty, 2004). Consider the following problem: You are presented with a sequence of gears (see Figure 7.1), and told that the leftmost gear is rotated in a clockwise direction. Your task is to predict the direction of movement of the rightmost gear. How do you solve this prediction problem? Psychological studies (e.g. Schwartz & Black 1999) suggest that people engage in mental simulation of the gear system. They tend to mentally rotate the first gear, and imagine how the second gear will rotate in response (it helps if you include some of the gear teeth!). This is continued in a piecemeal fashion until the final gear is rotated. After practice with this way of solving the problem, people often graduate to a more analytic solution, whereby they acquire the rule that adjacent gears will move in opposite directions. But for this task, and a host of other problems (e.g. pulley systems, water in glasses, etc.), people predominantly use mental simulation to answer questions of causal inference.

Hegarty (2004) has extracted several important principles from these studies:

(1) Simulation of complex physical systems is piecemeal rather than holistic, and occurs in the direction of causality (and time). For instance, to solve the gear problem, people simulated the gears sequentially, in a chain, starting from the initial cause and leading onto the final effect. In

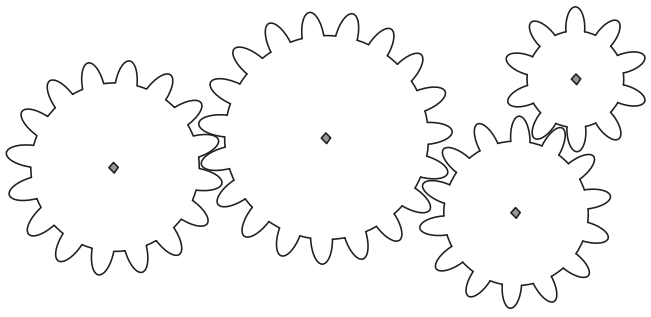

Fig. 7.1 System of interlocking gears.

another example involving pulley problems, Hegarty (1992) found that when people had to infer the movement of a pulley located in the middle of a causal chain, their eye fixations implied that they simulated causally prior pulleys that led to the movement of the pulley in question, but not pulleys that were causally downstream of this intermediate pulley.

These findings suggest that simulation does not operate on a complete or wholesale model of the physical set-up, but proceeds by selectively operating with smaller sub-components (i.e. individual causal links)[6]; it also suggests that simulations are constrained by the limitations of working memory.

(2) Simulation is not solely based on visual information, but can include non-visual information such as force and density. For example, people's mental simulations of the movements of liquids in a container are sensitive to the effects of gravity and the viscosity of the liquid (Schwartz, 1999). This suggests that mental simulation does not simply correspond to the manipulation of visual images, but can incorporate more abstract variables that explain a system's behaviour in the world.

(3) Simulations can include motor representations, especially when people are simulating their own (or other's) actions. Indeed there is now a rich literature on the role of motor representations in thinking (Jeannerod, 2006), and some sophisticated computational models of action that use internal models to predict sensory consequences of both actual and imagined actions (Wolpert, 2007).

(4) People use a variety of strategies to solve mechanical inference problems; these include mental simulation, but also rule-based strategies and analogical reasoning. These strategies are not mutually exclusive, and thus people might use a combination of strategies to solve a problem. As noted above, Schwartz & Black (1999) found that in the gear problems people progressed from using mental simulation to formulating and implementing a simple rule. Schwartz and Black speculate that mental simulation is best suited to novel problem situations, where people cannot draw on a ready-made set of formal rules.

In addition to these principles, Hegarty distinguishes between visual and spatial representation. The former represents visual aspects of things, such as colour and brightness, whereas the latter represents spatial relations and location and movement in space. Hegarty argues that mechanical reasoning predominantly depends on manipulations of spatial rather than visual images.

[6] Applied to the causal learning literature, this fits with suggestions made by Lagnado *et al.* (2007) and recent empirical work by Fernbach & Sloman (2008).

The work on mental simulation in mechanical reasoning has garnered strong empirical support (see Hegarty 2004, Nersessian, 2008). There remain open questions about the exact nature of simulation (e.g. how and in what respects do mental simulations track the real world causal processes), but the descriptive claim that people engage in mental simulation is generally well accepted. It also seems relatively straightforward to apply these ideas to the psychology of causal inference (for a related program see Wolff, 2007). The key idea would be that when people reason about causal systems they utilize mental models that represent objects, events or states of affairs, and reasoning and inference is carried out by mental simulation of these models. Moreover, these mental models admit of multifarious formats ranging from visual or spatial images, sensorimotor models to amodal representations.

7.3.2 Mental simulation of interventions

One of the most basic kinds of simulation involves the predictions of the effects of our own motor actions. In such cases mental simulation is likely to be tied quite closely to sensorimotor representations (e.g. forward models, Jeannerod, 2006, Wolpert, 2007), although these simulations can incorporate more abstract and non-visual elements too (e.g. gravity; friction etc). There is a natural progression to the simulation of other's actions (Brass and Heyes, 2005), and then to actions that need not involve an agent – e.g. natural causes such as the wind blowing; fire spreading etc. In these contexts the notions of cause and effect become less tied to agency and actions and their immediate effects. Note the parallel in the development of interventionist theories of causation in philosophy. These theories were initially tied to human agency (Collingwood, 1940), but subsequently developed in terms of potential manipulations without anthropomorphic connotations (Woodward, 2003). The evolution of mental causal simulation might have followed a similar course – at first tied to first-person agency and the immediate effects of actions, but progressing to simulations that need not involve agents, and incorporate more abstract causal variables.

One advantage of linking causal inference to simulation is that it can explain a variety of empirical findings which show that inference is enhanced with concrete materials and when spatial imagery/ visualization are supported (Barsalou, 1999; 2003). It also provides a ready explanation for situations where people are misled in their inferences. For instance, when the ease of mental simulation is mistakenly taken as an accurate guide to what actually happened. A compelling example of this is the power of direct witness testimony in legal cases (Heller, 2006). Vivid details given by an eyewitness about how the accused committed the crime greatly aid the jurors in imagining

978–0–19–957413–1 07-Mckay-Illari-c07-drv Mckay (Typeset by SPi, Chennai) 143 of 928 October 30, 2010 12:47
OUP UNCORRECTED PROOF – FIRST PROOF, 30/10/2010, SPi

this scenario, and facilitate the move from speculation to conviction. Indeed computer animations that attempt to reconstruct the visual aspects of a crime are increasingly popular in legal cases.

However, as well as showing how people's causal inferences might get distorted, the simulation account can also explain how they can make inferences in accord with normative models of causality. This is because mental simulation, by the very nature of being a simulation of the external causal system, will automatically observe the basic causal 'logic' of that external system. For example, simulating forwards from cause to effects naturally obeys the logic of intervention whereby an intervened-on variable propagates its effects downstream but not upstream (i.e. obeys 'do' surgery). For example, when I imagine myself intervening to turn the leftmost gear in Figure 7.1 I simulate the consequences of this action, such as the turning of the adjacent gear, but I do not typically simulate other possible causes of these effects, such as someone else turning the second gear instead. Predictive inference (inferring effects from causes) is thus relatively easy, because the system's behaviour is simulated forward (in time) from representations of causes to representations of effects. Diagnostic inference (inferring causes from effects) is more complex, because there might be various different possible causes of the effect in question, and hence a need to compare several simulations (for recent work that fits with the differences between diagnostic and predictive inference see Fernbach & Sloman 2009, also see Hagmayer & Waldmann, 2000). More complicated still are situations demanding both predictive and diagnostic inference. In such cases it is likely that people rely on piecemeal simulations of separate links (cf. Hegarty, 2004, Fernbach & Sloman, 2009). Indeed the simulation-based account of causal reasoning makes a range of testable predictions about the kinds of causal inference that people will find easy or hard, and the methods by which they might attempt to solve complex causal inferences.

One important question for this kind of approach is how it might be extended to causal inferences that do not involve directly or easily perceived causal systems? As well as making causal inferences in the physical or mechanical domain, we are able to reason about far more complex systems and about unobserved variables. Just think of the complexity in predicting the behaviour of your favourite soap-opera character. But the same issue arises with simpler cases too. Consider the inference that the low battery in the detector is causing the squeals. Presumably we need not know much about the actual physical processes that make the low battery cause the squeals. So what is involved in the hypothetical inference that removing the battery will stop the squeals? Do we still run simulations in such a situation?

One line of response to this issue is to note that mental simulation is a broad church, and accepts a plurality of possible representational formats – perceptual, motoric, amodal etc. Although its origins might lie in mental models that are closely tied to our immediate experiences with the

978–0–19–957413–1 07-Mckay-Illari-c07-drv Mckay (Typeset by SPi, Chennai) 144 of 928 October 30, 2010 12:47
OUP UNCORRECTED PROOF – FIRST PROOF, 30/10/2010, SPi

world (e.g. sensorimotor representations), these can become increasingly more abstract (e.g. spatial and map-like representations, cf. Tolman, 1948; Grush, 2004). Thus, although mental simulation can be accompanied by modal imagery (e.g. when you imagine a diver doing a somersault), visual imagery is not an essential part of simulation. It is possible to engage in mental simulation without explicit visual imagery. And it is also possible to use visual imagery to simulate a very abstract inference (e.g. imagining the economy taking a nose-dive). Indeed in cases of more abstract causal reasoning it is likely that representational schemes and models are created on the fly to solve specific inference problems, with different representations and mappings being used on different occasions. One day I might make predictions about the effects of the credit crunch by using a schematic model of a high-board diver, on another day I might prefer to use a model of a spreading fire, and at another time I might give up on imagery altogether. (This highlights the tight coupling between representation and inference, and the flexibility of our representational resources.)

The main point is that mental simulations are not restricted to sensorimotor models, but can incorporate a rich array of entities and processes. Indeed even the act of perception involves complex higher-level representations in terms of objects and their interrelations they bear. Thus mental models can be hierarchically structured, with components that can be recombined and reused (and are proposition-like in this respect).

Much of this is speculation ahead of empirical enquiry, but it is notable that all four principles advocated by Hegarty (2004) in the domain of mechanical reasoning seem to apply to the more general context of causal inference. Causal reasoning proceeds piecemeal (Fernbach & Sloman 2009; Lagnado *et al.* 2007; Hagmayer & Waldmann, 2000), it is not tied to visual representations, it takes advantage of motoric information (Jeannerod, 2006; Wolpert, 2007), and seems to admit of a variety of strategies ranging from image-based to abstract amodal simulation. The latter then paves the way for the formulation of general causal rules (see below). Suggestive evidence is also provided by the experiment described in an earlier section of this chapter (Lagnado & Loventoft-Jessen, in prep.). This experiment suggested that people were able to mentally represent possible causal models, and simulate possible interventions on these models. This included the ability to represent two different kinds of operation (disabling one slider and moving another), and draw appropriate inferences from this.

7.3.3 Is causal reasoning more basic than logical reasoning?

At this juncture it is useful to compare the proposed approach with the standard mental model theory advanced by Johnson-Laird and colleagues. The latter theory assumes that when people engage in deductive reasoning they use iconic but amodal representations. The theory then explains people's well-

978–0–19–957413–1 07-Mckay-Illari-c07-drv Mckay (Typeset by SPi, Chennai) 145 of 928 October 30, 2010 12:47
OUP UNCORRECTED PROOF – FIRST PROOF, 30/10/2010, SPi

known shortcomings in logical reasoning by adding principles that capture people's failures to represent the full set of possibilities (e.g. focus just on A & B when entertaining 'If A then B'). An alternative approach would be to accept that people can sometimes use amodal representations, but argue that the primary form of inference is causal – via mental simulation of these representations (which might include aspects that are modal), and that logical reasoning piggy-backs on this ability. This would explain many of the shortcomings in logical reasoning (e.g. influence of content and context), and also explain people's superior capability for causal reasoning. A similar argument might be made with respect to recent claims that logical reasoning is subserved by probabilistic reasoning (Oaksford & Chater, 2007). Here again the many shortcomings in people's explicit probabilistic reasoning might be explicable by their use of (causal) mental simulations in these situations (see Krynski & Tenenbaum, 2007, and Kahneman & Tversky, 1982, for related arguments). Of course the details of such an argument need to be spelled out in much more detail, and empirical studies need to be designed in this light. But it seems a suggestive possibility.

Moreover, it yields a simple account for how people can slowly acquire mastery of logical and probabilistic reasoning. They gradually capitalize on their ability to manipulate amodal representations, and integrate this with a more sentence-like symbolic language presented to them while they are learning. This process resembles the observed transition from simulating gears to learning a formal rule (Schwartz & Black, 1999). However, the role of some kind of imagery is probably never lost, and can persist even in rarefied domain of scientific inference (Hadamard 1954).

In short, the speculative claim here is that people have a core capability for *causal* reasoning via mental simulation, and deductive and inductive reasoning builds on this foundation. This shift of perspective might explain the characteristic biases and shortcomings in lay people's logical and probabilistic reasoning.

7.4 Conclusions

I have argued for the interplay of causal learning and reasoning, the multiplicity of sources of evidence for causal relations, and the role of mental simulation in causal inference. These three strands are themselves intertwined. Learning and reasoning utilize the same kinds of representations and mechanisms: they both rely on mental models, and in both cases inference depends on the simulation of these models. The fact that these mental models admit of multifarious formats (e.g. spatial, perceptual, sensorimotor) reflects the rich causal information available from the world and our interactions with it. Nevertheless, our ability to construct evermore abstract amodal representations,

catalysed by the invention of external representational forms such as diagrams and language, enables us to draw causal inferences that take us beyond the surface of our perceptions.

References

Ahn, W. & Kalish, C. (2000). The role of covariation vs. mechanism information in causal attribution. In R. Wilson, & F. Keil (eds.) *Cognition and Explanation,* Cambridge, MA: MIT Press.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences,* 22, 577–660.

Barsalou, L.W., Simmons, W. K., Barbey, A. K. & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-speci?c systems. *Trends in Cognitive Science.*, 7, 84–91.

Braine, M. & O'Brien, D. (1998). *Mental Logic*. Mahwah, NJ: Lawrence Erlbaum Associates.

Brass, M. & Heyes, C. M. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Science,* 9, 489–495.

Bower, G. H. (1970). Organizational factors in memory. *Cognitive Psychology,* 1, 18–46.

Cartwright, N. (2007). *Hunting Causes and Using Them: Approaches in Philosophy and Economics,* Cambridge: Cambridge University Press.

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

Collingwood, R. (1940). *An Essay on Metaphysics*. Oxford: Clarendon Press.

Cowan, N. (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.

Craik, K. (1952). *The Nature of Explanation. Cambridge University Press, Cambridge, UK.*

Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99, 3–19.

Evans, J. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin,* 128, 978–96.

Fernbach, P. M. & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology: Learning, Memory, and Cognition. 35 (3), 678–693.*

Glymour, C. (2007). Statistical jokes and social effects: Intervention and invariance in causal relations. In Gopnik, A., & Schultz, L. (eds.), *Causal learning: Psychology, Philosophy, and Computation,* pp. 294–300. Oxford: Oxford University Press.

Goldvarg, Y. Johnson-Laird, P.N. (2001) Naïve causality: a mental model theory of causal meaning and reasoning. *Cognitive Science,* 25, 565–610.

Gopnik, A. & Schultz, L. (2007). *Causal Learning: Psychology, Philosophy, and Computation*. Oxford: Oxford University Press.

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

Griffiths, T.L. & Tenenbaum, J.B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, 51, 334–384.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27, 377–396.

Hadamard, J. (1954). *The Psychology of Invention in the Mathematical Field*. New York: Dover.

Hagmayer, Y., & Waldmann, M. R. (2000). Simulating causal models: The way to structural sensitivity. In L. Gleitman & A. Joshi (Eds.), *Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Hall, G. (2002). Associative structures in Pavlovian and instrumental conditioning. In C.R. Gallistel (ed.), *Stevens' Handbook of Experimental Psychology*, 3rd edition, Vol. 3, pp. 1–45. New York: John Wiley.

Hastie, R. & Pennington, N. (2000). Explanation-based decision making. In T. Connolly, H. R. Arkes and K. R. Hammond (eds): *Judgment and Decision Making: An Interdisciplinary Reader* (2nd ed.). pp. 212–28. Cambridge University Press.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: a new hypothesis. *Trends in Cognitive Sciences,* 11 (6), 236–242.

Hegarty, M. (2004). Mechanical reasoning as mental simulation. *Trends in Cognitive Science,* 8, 280–285.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 18, 1084–1102.

Heller, K. J. (2006). The cognitive psychology of circumstantial evidence. *Michigan Law Review,* 105, 241–305.

Inhelder, B. and Piaget, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. New York: Basic Books.

Jeannerod, M. (2006) *Motor Cognition: What Actions Tell the Self*. Oxford: Oxford University Press.

Johnson-Laird, P.N. (1983) *Mental Models*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. (2006). *How We Reason*. Oxford: Oxford University Press.

Kahneman, D. & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (eds.), *Judgment under Uncertainty: Heuristics and Biases* (pp. 201–208). New York: Cambridge University Press.

Krynski, T. R. & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General,* 136, 430–450.

Lagnado, D.A. (2009). A causal framework for learning and reasoning. *Behavioral and Brain Sciences,* 32, (2), 211–212.

Lagnado, D.A. (2010). Thinking about evidence. To appear in Dawid, P, Twining, W., Vasilaki, M. eds. *Evidence, Inference and Enquiry*. British Academy/OUP. (In Press).

Lagnado, D.A. & Sloman, S.A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory & Cognition,* 30, 856–876.

Lagnado, D.A. & Sloman, S.A. (2002). Learning causal structure. In W. Gray & C. D. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, Erlbaum.

Lagnado, D.A., Waldmann, M.R., Hagmayer, Y., & Sloman, S.A. (2007). Beyond covariation: Cues to causal structure. In Gopnik, A., & Schultz, L. (eds.), *Causal learning: Psychology, Philosophy, and Computation,* pp. 154–172. Oxford: Oxford University Press.

Lagnado, D.A. & Harvey, N. (2008). The impact of discredited evidence. *Psychonomic Bulletin and Review*, 15, 1166–1173.

Lagnado, D.A. & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory & Cognition,* 32, 451–460.

Lagnado, D. A. & Loventoft-Jessen, J. (in prep.). Learning causal models through multiple interventions.

Meder, B., Hagmayer, Y., & Waldmann, M. R. (2008). Inferring interventional predictions from observational learning data. *Psychonomic Bulletin & Review,* 15, 75–80.

Michotte, A. (1954/1963). *The Perception of Causality*. London: Methuen.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 63, 81–97.

Nersessian, N. J. (2008). *Creating Scientific Concepts*. Cambridge, MA: MIT Press.

Oaksford, M. & Chater, N. (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford: Oxford University Press.

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

Rips, L. J. (1983). Cognitive processes in prepositional reasoning. *Psychological Review,* 90, 38–71.

Shanks, D. R. (2004). Judging covariation and causation. In D. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making*. Oxford: Blackwell.

Shanks, D. R. (1995). *The Psychology of Associative Learning*. Cambridge: Cambridge University Press.

Schlottmann, A. (2001). Perception versus knowledge of cause-and-effect in children: When seeing is believing. *Current Directions in Psychological Science*, 10, 111–115.

Scholl, B. J. (2005). Innateness and (Bayesian) visual perception: Reconciling nativism and development. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate Mind: Structure and Contents* pp. 34–52. Oxford: Oxford University Press.

Schwartz, D.L. (1999). Physical imagery: kinematic vs. dynamic models. *Cognitive Psychology*. 38, 433–464.

Schwartz, D.L. & Black, T. (1999). Inferences through imagined actions: knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory & Cognition,* 25, 116–136.

Shanks, D. R. & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* Vol. 21, pp. 229–261. San Diego, CA: Academic Press.

Sloman, S. A. (2005). *Causal Models; How People Think About the World and its Alternatives*. New York: Oxford University Press.

Sloman, S. A., Barbey, A. K. & Hotaling, J. (2009). A causal model theory of the meaning of cause, enable, and prevent. *Cognitive Science,* 33, 21–50.

Sloman, S. A. & Lagnado, D. A. (2005). Do we do? *Cognitive Science,*. 29, 5–39.

Spirtes, P., Glymour, C., & Schienes, R. (1993). *Causation, Prediction, and Search*. New York: Springer-Verlag.

Steel, D. (2008). *Across the Boundaries: Extrapolation in Biology and Social Science*. New York: Oxford University Press.

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

Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.

Tenenbaum, J. B., & Griffiths, T. L. (2003). Theory-based causal inference. *Advances in Neural Information Processing Systems 15* pp. 35–42. Cambridge, MA: MIT Press.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208.

Wagemans, J., van Lier, R. & Scholl, B. J. (2006). Introduction to Michotte's heritage in perception and cognition research. *Acta Psychologica,* 123, 1–19

Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The Psychology of Learning and Motivation* Vol. 34, pp. 47–88. San Diego, CA: Academic Press.

Waldmann, M. R. & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: asymmetries in cue competition. *Journal of Experimental Psychology, General,* 121(2), 222–236.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Motivation, and Cognition*, 31, 216–227.

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 (6), 307–311.

Wason, P. C. (1983). Realism and rationality in the selection task. In J. St. B. T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches*. London: Routledge & Kegan

Williamson, J. (2005). *Bayesian Nets and Causality: Philosophical and Computational Foundations*. Oxford: Oxford University Press.

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

Wolpert, D. M. (2007). Probabilistic models in human sensorimotor control. *Human Movement Science*, 26, 511–524.

Woodward, J. (2003). *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.