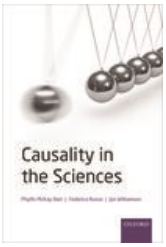


Causality in the Sciences

Phyllis McKay Illari (ed.) et al.

https://doi.org/10.1093/acprof:oso/9780199574131.001.0001

Published: 2011 **Online ISBN:** 9780191728921 **Print ISBN:** 9780199574131

CHAPTER

8 8 When and how do people reason about unobserved causes?

Benjamin Rottman, Woo-kyoung Ahn, Christian Luhmann

https://doi.org/10.1093/acprof:oso/9780199574131.003.0008 Pages 150–183

Published: March 2011

Abstract

Assumptions and beliefs about unobserved causes are critical for inferring causal relationships from observed correlations. For example, an unobserved factor can influence two observed variables, creating a spurious relationship. Or an observed cause may interact with unobserved factors to produce an effect, in which case the contingency between the observed cause and effect cannot be taken at face value to infer causality. This chapter reviews evidence that three types of situations lead people to infer unobserved causes: after observing single events that occur in the absence of any precipitating causal event, after observing a systematic pattern among events that cannot be explained by observed causes, and after observing a previously stable causal relationship change. In all three scenarios people make sophisticated inferences about unobserved causes to explain the observed data. This chapter discusses working memory as a requirement for reasoning about unobserved causes and briefly discuss implications for models of human causal reasoning.

Keywords: unobserved causes, causal learning, causal inference

Subject: Logic

Collection: Oxford Scholarship Online

Abstract

Assumptions and beliefs about unobserved causes are critical for inferring causal relationships from observed correlations. For example, an unobserved factor can influence two observed variables, creating a spurious relationship. Or an observed cause may interact with unobserved factors to produce an effect, in which case the contingency between the observed cause and effect cannot be taken at face value to infer causality. We review evidence that three types of situations lead people to infer unobserved causes: after observing single events that occur in the absence of any precipitating causal event, after observing a systematic pattern among events that cannot be explained by observed causes, and after observing a previously stable causal relationship change. In all three scenarios people make sophisticated inferences about unobserved causes to explain the

> observed data. We also discuss working memory as a requirement for reasoning about unobserved causes and briefly discuss implications for models of human causal reasoning.

An observed correlation between two events does not imply a direct causal relationship between them. One reason that is particularly important in developing theories of human causal learning is that unobserved or unattended cause(s) may account for all or part of the observed correlations.

For instance, an article published in Nature reported that young children who sleep with a nightlight are much more likely to develop myopia later in life (Quinn, Shin, Maguire, & Stone, 1999). This was interpreted as implying a causal relationship by the popular press. For instance, CNN reported, even low levels of light can penetrate the eyelids during sleep, keeping the eyes working when they should be at rest. Taking precautions during infancy, when eyes are developing at a rapid pace, may ward off vision trouble later in life (CNN, May 13, 1999). A later study, however, suggested that a common cause is responsible for this correlation; myopic parents are more likely to leave a light on for children, and myopic parents are more likely to have myopic children (Gwiazda, Ong, Held, & Thorn, 2000).

While the above example illustrates how a positive correlation between two variables does not imply that
p. 151 one causes the other, the opposite can happen ↳ as well; we observe no correlation between two events when, in fact, there is a causal relationship between them. For example, a recent study demonstrated that pollution and daily temperature range are positively associated in the summer, but negatively associated in the winter (Gong, Guo, & Ho, 2006). Prior to learning that season plays a causal role, it would appear as if there is no relationship between pollution and temperature range because there is no correlation, even though there is an important relationship. Restated, there was a period of time during which an unknown variable (season) obscured the causal relationship between two observed variables, and the researchers had to learn about this interaction.

Considering these examples, it should be obvious that assumptions and beliefs about unobserved causes are vital in inferring causal relationships from observed correlations. In some sense, it is remarkable that we can make any valid causal inferences from observed correlations alone. There can be any number of unobserved causes at play, and people cannot possibly reason through all possible combinations whenever they make causal inferences.

This paper examines laypeoples' inferences about unobserved causes. We will first elaborate on the problems involving unobserved causes. Then, we will argue that people actually perform fairly sophisticated reasoning about unobserved causes, and that such reasoning is engaged due to a certain set of assumptions that they hold about the world. We also review psychological studies supporting our argument.

Problems with reasoning about unobserved causes

Consider a simple causal reasoning scenario involving a light switch and a light. Suppose you go into a room for the first time, and you observe the light (i.e. on or off; 1 or 0, respectively in Table 8.1 under Light) across eight consecutive trials when the switch is up or down (1 or 0 respectively in Table 8.1 under Switch). One possible causal interpretation is that there is no causal relationship between the observed switch and the light, and there is an unobserved switch that is *entirely responsible* for the light's behaviour (see Table 8.1, ‘Entirely Responsible’ column).

Table 8.1 Light switch example

Trials	Observed events		Different types of possible unobserved switches			
	Switch	Light	Entirely Responsible	Biconditional	2 Out of 3	Always Present
1	0	0	0	1	00	1
2	1	0	0	0	00	1
3	1	1	1	1	10	1
4	0	1	1	0	11	1
5	0	0	0	1	10	1
6	1	0	0	0	00	1
7	1	1	1	1	11	1
8	0	1	1	0	11	1

But, there are many other equally plausible possibilities in which the observed switch is causally responsible for the effect in combination with another unobserved switch (see Table 8.1). One is that the observed switch interacts with another switch through a *biconditional* interaction such that the light turns on only when the two switches are either both up or both down. Yet another possibility is that there are two unobserved switches, and at least *two out of three* of these switches must be up to make the light turn on. Depending on whether one believes in the biconditional interaction or two out of three unobserved switches, one's future intervention to make the light go on would change (e.g. if it is two unobserved switches, keeping the switch
p. 152 up would maximize the time the light is on, but for a biconditional ↳ case, flipping the switch whenever the light goes off would likely maximize the time the light is on).

The point of this example is to illustrate that there are so many possible ways that unobserved causes could interact that it would be impossible for people to consider all of these configurations. Does this mean that people do not spontaneously reason about unobserved causes? The answer must be no, given the obvious fact that people do make causal inferences based on correlations, and they must make (or act as if they make) some assumptions about unobserved causes in order to do so. (For instance, inferring that X causes Y based on a positive correlation between X and Y requires assuming that there is no unobserved, confounding variable.) The important question, then, is what assumptions and inferences people make about unobserved causes, and what triggers inferences about unobserved causes given that people cannot always consider all possible unobserved causes? The current chapter reviews studies from our labs that provide some answers to these questions.

In the following sections, we first briefly review how existing models of causal learning handle unobserved causes. Then we argue that people hold assumptions that trigger *specific* inferences about unobserved causes. We claim that people believe that (i) an event must be caused by another event (causal determinism), (ii) any systematic pattern or regularity among events must be causally determined, and (iii) causal relations stay stable across different times and contexts. When causal determinism is violated, when a systematic pattern is not explained by observed causes, or when causal relations are not stable, we argue that people infer an unobserved cause to explain the apparent violation of the assumption. Then we present experimental results suggesting that people do spontaneously make such inferences about unobserved causes, and describe how such inferences further influence the causal inferences people draw from observed correlations. Finally, we will discuss one cognitive requirement for reasoning about unobserved causes.

p. 153

8.1 Unobserved causes in models of human causal learning

Many models have been developed to explain how people learn the causal strength of a particular cause and effect relationship. Luhmann and Ahn (2007) and Hagmayer and Waldmann (2007) have provided detailed reviews of how these models handle unobserved causes. Here we provide a brief summary.

One class of models makes no assumptions about unobserved causes, and thus makes no inferences about unobserved causes. For example, AP (Jenkins & Ward, 1965), an associative measure, estimates causal strength as the difference in probability of the effect (E) being present when the cause (C) is present vs. absent: $P(E|C) - P(E| \sim C)$. Though ΔP is a very intuitive way of calculating the influence of C on E, it runs into a critical problem; people are more sensitive to certain types of evidence such as when both C and E are present and are less sensitive to other types of evidence such as when both are absent. Many subsequent descriptive models have tried to capture this phenomenon by differentially weighting the evidence (e.g. Arkes & Harkness, 1983; Downing, Steinberg, & Ross, 1985; Einhorn & Hogarth, 1986; Nisbett & Ross, 1980; Schustack & Sternberg, 1981; Shaklee and Tucker, 1980). However, these approaches did not provide a theoretical explanation for the phenomenon.

Cheng (1997), Novick and Cheng (2004), see also Pearl (2000), provided a parsimonious theoretical explanation for this phenomenon by appealing to unobserved causes. Cheng argued that differential weighting of evidence is a normative result of accounting for ceiling effects, when an unobserved cause frequently produces the effect (see Section 8.4.1 for a discussion). However, Cheng's model requires a number of assumptions. Specifically, for a generative observed cause, unobserved causes are assumed to interact in a noisy-or fashion with observed causes, to be generative, not inhibitory, and to be independent from observed causes. These very strict assumptions limit the applicability of the model and it is not entirely clear whether people actually make these assumptions (Luhmann & Ahn, 2007; Hagmayer & Waldmann, 2007; White, 2005, 2009).

A very different approach to unobserved causes makes the straightforward assumption that all unobserved causes, taken as a whole, are *present* across all learning trials. For example, the Rescorla-Wagner model (Rescorla & Wagner, 1972; Dickinson, 1984) includes a background context node that can be viewed as an aggregation of all unobserved causes. When an effect occurs without the observed cause, this node gains associative strength, which can be used as an estimate of the causal strength of an unobserved cause. However, it is easy to see that the consequence of this assumption would be quite unsatisfactory for reasoners. For example, in Table 8.1, last column, an unobserved cause present on every trial would be completely unable to explain the light's behaviour; neither the observed switch nor the unobserved cause
p. 154 correlates ↳ with the status of the light and thus the only unsatisfactory conclusion is that the light was acting randomly without any cause.

Some recent models have attempted to explain peoples' sophisticated reasoning about unobserved causes including (i) inferring whether an unobserved cause is present or absent on a particular trial, and (ii) inferring the causal strength of an unobserved cause. For example, if an effect is observed without an observed cause, one would likely infer that an unobserved cause is responsible. Furthermore, given that an observed cause is present, one would more likely infer that an unobserved cause is also present if the effect is present rather than absent (see Hagmayer & Waldmann, 2007, for a detailed explanation of these examples). One new model, BUCKLE (Luhmann & Ahn, 2007), has been developed specifically for these types of inferences. BUCKLE is explained in the Section 8.2.

In sum, few models have been developed to account for reasoning about unobserved causes, though there have been some recent attempts to explain how people learn about the presence and causal strength of an unobserved cause. In the next section, we provide further evidence of reasoning about unobserved causes that a more comprehensive model should account for.

8.2 Causal determinism about individual events

One of the more primitive assumptions that lay reasoners appear to make is causal determinism, that every event has a cause and that events cannot occur in the absence of any precipitating causal event.[1] This assumption of causal determinism is captured in the causal principle from ancient philosophy; 'nothing happens without a cause' ('*nihil fit sine causa*' Audi, 1995). For someone who believes in causal determinism, events with no apparent cause should suggest the existence of hidden causes. Much of the empirical work suggesting that people believe in causal determinism has investigated children's beliefs about agency and magic, which is reviewed below.

8.2.1 Children

Children's beliefs about agency

A major question in developmental psychology pertains to children's beliefs about agency, the idea that there are entities with free will (e.g. humans and animals) that are primary sources of causal influence. For example, the motion of animate agents may be assumed to be generated internally and to not require
p. 155 further explanation (e.g. Wegner, 2002; Leising, Wong, Waldmann, & ↳ Blaisdell, 2008). In contrast, the motion of non-agents (e.g. billiard balls) must be explained by referring to external causal forces. When an inanimate object assumed not to have self-agency appears to move on its own (e.g. a baseball moving like a bird rather than in an arc), this violation of determinism should be surprising.

Saxe, Tenenbaum, & Carey (2005), see also Saxe, Tzelnic, & Carey (2007), tested this reasoning with infants in the following way. They had infants repeatedly observe a beanbag (a non-agent) flying through the air from one side of a small stage to the other. (Studies with infants often use a 'habituation' phase, during which the infant becomes accustomed to seeing the same event and stops paying attention to the event. In a later phase, if infants show increased interest, this is taken to imply 'surprise.') After an infant was habituated to this event, he/she was presented with this same event (the beanbag flying across the stage) followed by a human hand entering from one side of the stage, either the side from which the beanbag was launched, or the opposite side. The infants were more 'surprised' (spent more time looking at the hand) when the hand entered from the opposite side of the stage as the beanbag. The reasoning is that when the hand entered from the same side as the beanbag, the infants could reason backwards that the hand had been behind the stage all along and could have thrown the beanbag. However, when the hand entered from the opposite side as the beanbag, there is no cause of the beanbag's motion — the beanbag cannot propel itself and the hand was on the opposite side. Critically, this result disappeared entirely when the beanbag was replaced with a puppet (an agent) that appeared to propel itself across the stage without requiring another agent. In sum, the infants were 'surprised' only when an inanimate object appeared to move itself, a violation of causal determinism, but they were not surprised when an animate object, assumed to have self-agency, moved itself.

Children's beliefs about magic

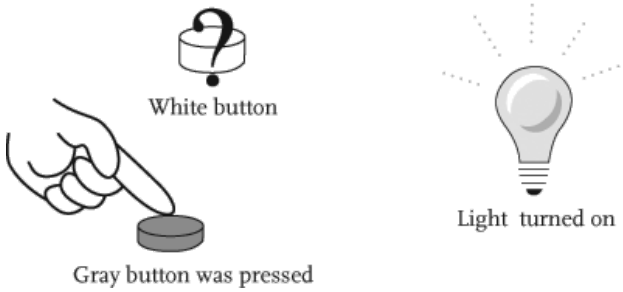
Other developmental work has examined circumstances that evoke magical explanations from children (see Woolley, 1997 for a review). For example, Phelps & Woolley (1994) presented children (ages 4—8) with several real- world objects and asked them about their operation. Children were shown two objects that were, unbeknownst to the children, magnets of opposing polarity. The children were first asked to make a prediction (e.g. whether one object could move the other without touching it) and then to provide an explanation once a surprising event occurred (e.g. after one object pushed the other without touching it). This study revealed that if a child could not explain the event with a physical explanation, he/she tended to appeal to magic or 'tricks' (both of which refer to hidden causes). Thus, children's reliance on magic for explanation, an inference to the ultimate hidden cause, appears to be strongly driven by events that violate causal determinism.

p. 156

8.2.2 Adults

Luhmann & Ahn (2007) have recently conducted a series of experiments to explore beliefs about causal determinism in adults. The study was designed to investigate whether adults make significant inferences to hidden causes and what, if any, influence on behavior such simple inferences might have. The study used a typical causal learning task in which subjects were asked to learn about a pair of potential causes (gray and white buttons) and their influence on a single effect (light turning on). Participants observed the presence/absence of the different events in a trial-by-trial manner. On each trial, learners observed the presence/absence of one cause (whether the gray button in Figure 8.1 was pressed or not) and the presence/absence of the effect (whether the light was on/off). Unlike typical causal learning experiments, one of the two causes in our study was 'hidden' from subjects (the white button in Figure 8.1). No information was ever provided to the participants about the presence/absence of the second, hidden cause. After they completed the trial sequence, participants evaluated the strength of the causal relationship between the observed cause and the effect and the strength of the relationship between the hidden cause and the effect.

Fig. 8.1



Sample unexplained effect trial used in Luhmann & Ahn (2007).

To determine whether learners made notable inferences about the hidden cause, we manipulated whether or not the different sequences included trials that violated causal determinism; trials in which the effect was present but the observed cause was absent (the light was on, but the button was not pressed; Figure 8.1). We call these trials unexplained effects, as effects are present in the absence of any observed causes. As shown in Table 8.2, the 'unnecessary' and 'zero' conditions included unexplained effects, whereas the 'perfect' and 'insufficient' conditions did not. Our results demonstrated that sequences that included unexplained effects, or violations of causal determinism, led subjects to believe that the hidden cause was a stronger

p. 157 (generative) cause than sequences that did not include unexplained effects (last column of Table 8.2). ↳

Table 8.2 Summary of conditions and mean probability and causal strength estimates of unobserved cause in Luhmann & Ahn (2007, Experiment 3)*.

Unovserved Effects	Conditions	Observed frequencies of four trial types	Average trial-by-trial likelihood judgments of hidden cause being present (top) vs. absent (bottom) on each trial type	Average Causal Strength Estimates of Hidden Cause
Present	Unnecessary	E ~E O 7 0 ~O 7 7	E ~E O 5.8 . 5.8 ~O 7.3 2.1 4.8 6.7 2.1	70.8 (6.44)
	Zero	E ~E O 7 7 ~O 7 7	E ~E O 3.4 3.4 4.4 ~O 2.8 1.8 4.8 6.6 2.6	74.4 (5.35)
Absent	Perfect	E ~E O 7 0 ~O 0 7	E ~E O 4.1 . 4.1 ~O . 2.1 3.2 4.1 2.1	32.7 (6.83)
	Insufficient	E ~E O 7 7 ~O 0 7	E ~E O 5.3 3.8 4.5 ~O . 2.7 2.7 5.3 3.4	50.8 (7.56)

* *Note*: *O* is the observed cause, and *E* is the effect. ~ represents the absence of an event. Unexplained effect (~*OE*) trials are shown in bold. Standard errors are in parentheses.

Because violations of causal determinism were prima facie evidence for the operation of an unobserved, generative cause, we suggested that subjects were using these specific occasions as the basis for their causal strength judgments of the hidden cause.

To validate our explanation, we asked learners on each trial to judge how likely the hidden cause was present using a scale that ranged from 0 (definitely absent) to 10 (definitely present, see the fourth column, Table 8.2). These probability judgments allowed us to directly measure learners' beliefs about the hidden cause on all four types of trials. As expected, learners believed that the hidden cause was likely present when causal determinism was violated (~*OE* trials in bold in the fourth column in Table 8.2). In fact, learners believed that the hidden cause was more likely to be present on these occasions than on any other type of trial. Thus, similar to infants, violations of causal determinism lead adults to infer hidden causes.

The finding that people infer unobserved causes during unexplained effects may seem fairly intuitive. However, this experiment allowed us to uncover additional, potentially less intuitive, and more sophisticated inferences about hidden causes.

One demonstration of sophisticated reasoning about unobserved causes is that participants' real-time judgments about the presence/absence of the unobserved cause explain their judgments of the causal
p. 158 strength of the ↳ unobserved cause. To demonstrate this, we computed ΔP, a measure of covariation, between the unobserved cause and the effect (i.e. $P(E|U)-P(E|\sim U)$) based on participants' average probability judgments of the presence of the unobserved cause shown in Table 8.2 (converted to probabilities that ranged from 0 to 1); they were 0.40, 0.40, 0.10, and 0.19, respectively, for the four conditions in Table 8.2.[2] Impressively, these virtual covariations correlate with the subsequent judgments of the causal strength of the unobserved cause (shown in the last column of Table 8.2). It should be stressed that these were 'virtual' covariation in the sense that it only existed in the heads of the learners. No actual covariation existed because one of the two potentially covarying events was hidden. Furthermore, subjects were never asked to estimate the overall covariations between these two events; we computed them over participants' probability judgments of the presence of the unobserved cause. That is, when the virtual ΔP between the hidden cause and the effect was higher, subjects' causal strength estimate of the hidden cause

was higher, and vice versa, as if the 'virtual' covariation we computed had been directly observed by learners.

Even more subtly, each subject's idiosyncratic beliefs about hidden cause- effect covariation could be used to estimate his/her own inferences of the causal strength of the unobserved cause. Some subjects believed in strong covariation between the hidden cause and the effect. Other subjects' probability judgments showed weaker covariation. Remarkably, the individual differences in this virtual covariation measure (i.e. ΔP between the unobserved cause and the effect) significantly predicted subjects' subsequent judgments of the causal strength of the hidden cause in each of the four conditions: *rs* = 0.48, 0.43, 0.59, and 0.52, respectively (all *ps* ⟨ 0.05, Luhmann & Ahn 2007, unpublished analyses). Those subjects whose inferences implied strong hidden cause-effect covariation judged the causal relationship to be stronger than those subjects whose inferences implied weak covariation. This pattern of beliefs suggests particularly elaborate reasoning about hidden causes.

These data could be also used to evaluate some of the theoretical claims about hidden causes. For example, as mentioned above, prominent theories of causal inference (e.g. Cheng, 1997) require that hidden causes occur independently of observed causes; that is, the likelihood of a hidden cause, *U*, in the presence of an observed cause, *O*, is the same as the likelihood of *U* in the absence of *O*, $P(U|O) = P(U|{\sim}O)$. In contrast, according to subjects' probability judgments shown in Table 8.2, this requirement was violated in the
p. 159 majority of situations we tested. The hidden cause was judged to be ↳ more likely when the observed cause was present and less likely when the observed cause was absent (i.e. $P(U|O) \rangle P(U|{\sim}O)$), as illustrated by the marginal means of *O* and ~*O* in the fourth column of Table 8.2. Nonetheless, subjects were uniformly willing to estimate the strength of both the hidden and observed cause. This suggests that people might not believe that independence of hidden causes is a requirement for valid causal inference (Luhmann & Ahn, 2005; see also Hagmayer & Waldmann, 2007).

These data also provide insight into the conditions under which people infer unobserved causes to be generative or inhibitory. In the previous studies, the unobserved cause was always judged to be generative. However, Schulz and Sommerville (2006) demonstrated that four-year-olds sometimes infer preventative hidden causes. In their study, children were presented with a cause that produced an effect four times. They then observed eight trials when the cause unreliably produced the effect (sometimes the effect was present when the cause was present, sometimes the effect was absent when the cause was present). Finally, the children were shown a button box that the experimenter had hidden during the cause—effect sequence. When asked to prevent the effect, children pressed the previously hidden button, indicating that they thought it was preventative. To summarize, Schulz and Sommerville found that instances when a cause is present but the effect is absent (*O*~*E* observations) lead children to infer an inhibitory cause, but Luhmann and Ahn found that adults inferred a generative cause.

Luhmann and Ahn (2007) reasoned that *O*~*E* observations could be interpreted in multiple ways. For instance, *O*~*E* may occur (a) because an unobserved cause prevented the effect from happening or (b) because the observed cause is not entirely sufficient to bring about the effect. Thus, if a learner believes that the observed cause is weak, then the learner does not have to infer that the unobserved cause is inhibitory in order to account for *O*~*E* observations. Indeed, in conditions with *O*~*E* observations in the experiment described above (e.g. the Insufficient and Zero conditions), learners believed that the observed cause was relatively weak and the hidden cause was relatively strong and generative (e.g. note the relatively high causal judgment of the unobserved cause in the Insufficient cause). However, if people already believe that the observed cause is strongly generative, they might infer an unobserved inhibitory cause to explain *O*~*E* evidence.

To reconcile our findings with those of Schulz and Sommerville (2006), we designed an experiment that provided pre-training to learners. This pre- training was designed to convince learners that the observed

cause was, on its own, a sufficient cause of the effect. Once this pre-training was complete, we then presented the same Insufficient condition we used in the experiment described above. In light of the pre-training, learners' judgments indicated that they believed that the hidden cause was preventative. This
p. 160 result suggests an important difference between *O*~*E* and ~*O E* . Observations of ~*O E* are ↳ violations of causal determinism and require inferring a hidden generative cause. In contrast, observations of *O*~*E* are somewhat more ambiguous. If learners entertain the possibility that the observed cause has no causal influence at all, or if they allow for the possibility that the observed cause produces its effect unreliably, then there is no need to appeal to hidden causes at all. Alternatively, if learners believe that the observed cause reliably produces its effect (e.g. the children in Schulz and Sommerville (2006) experiment or the adults after our pre-training), then observations of *O*~*E* suggest the operation of hidden, preventative causes.

8.2.3 BUCKLE: A model of unobserved cause learning

Because existing theories of causal inference were unable to account for these results, Luhmann and Ahn (2007) proposed an alternative account, instantiated as a computational model called BUCKLE. The basic operation of BUCKLE involves (1) making inferences about the presence or absence of hidden causes (via Bayesian inference) and (2) then adjusting beliefs about the causal strength of all causes (both hidden and observed). These two steps are performed on each trial.

In Step 1, BUCKLE makes use of four pieces of information to estimate the probability that *U* is present: whether *O* and *E* are present or absent and the causal strengths of *O* and *U* that were calculated on the previous trial. Suppose that the learner observes an *O E* trial. If the learner believes that *O* is strong, then there is little reason to posit that *U* is present because *O* could have produced *E*. The stronger the causal strength of *O*, the less likely that *U* is to be present. However, if *O* is weak, then *U* is needed to explain the presence of *E*. In fact, if *O* has zero causal strength, then *U must* be present, otherwise there is no way to explain the presence of *E*. Finally, suppose that the learner observes an ~*O E* trial. Like the previous case, *U* also must be present because *O* is not present to produce *E*. These last two examples show how BUCKLE embodies the assumption of causal determinism; if *O* is unable to produce *E* because it is absent or has zero causal strength, then *U* must have been present and must have caused it. This sort of reasoning drives BUCKLE's inference about the presence of *U* on a given trial.

In Step 2, BUCKLE updates its estimate of the causal strengths of *O* and *U*. To do so, BUCKLE first predicts whether *E* should be present or not based on its current causal strengths of *O* and *U*, knowledge about the presence/absence of *O*, and its inference about the presence of *U* (from Step 1). Intuitively, this prediction is made the following way. Suppose we know that *O* is absent. Then, the only way *E* could be present is by *U*. Thus, the probability that *E* would be present is the causal strength of *U* multiplied by the probability that *U* is present. Alternatively, if *O* is present, the probability of *E* being present is increased if *O* is strong, if *U* is
p. 161 strong, and if *U* is likely present. Once BUCKLE has made its prediction about *E*, BUCKLE calculates ↳ the difference between this prediction and knowledge about the actual presence/absence of *E*. This step is very similar to the Rescorla and Wagner (1972) model. If BUCKLE under-predicts *E*, then the causal strengths of the present causes are increased. If BUCKLE over-predicts *E*, they are decreased.

BUCKLE is capable of accounting for the patterns of inferences described above, both the trial-by-trial judgments of the probability of an unobserved cause being present and the causal strength judgments. Additionally, BUCKLE explains how inferences that people make about unobserved causes interact with their inferences about observed causes such as order effects (Luhmann & Ahn, 2007).

In summary, these studies have shown that when an event occurs that cannot be explained by an observed cause, people infer an unobserved cause to explain the event. These studies have so far focused on

inferences that people make about single events. In the next section, we will discuss inferences people make about unobserved causes from patterns of events.

8.3 Causal determinism about systematic patterns among events

We argue that when people perceive a pattern in a sequence of events, they are reluctant to treat it as purely accidental, and instead they infer that the pattern was planned or produced through a causal mechanism. A classic example involves the pattern of bombs dropped on London by the Nazis during World War II (Gilovich, 1991, pp. 19–21; Hastie & Dawes, 2001, pp. 160–161). Even though the locations of the bombings have since been shown to be statistically random, many British citizens thought they saw clusters of bombings, and consequently inferred that German bombers deliberately avoided locations where German spies lived, creating the perceived clustering. This is a perfect example of how people infer an unobserved cause to explain an observed pattern (even though the pattern is statistically absent).

Perhaps the most basic types of patterns from which people infer a causal mechanism are those studied in introductory statistics courses: differences between the mean scores of two groups and correlations between two variables. There exists extensive literature about causal learning of this sort (e.g. Cheng, 1997; Jenkins & Ward, 1965). Here we focus on other patterns from which people are likely to infer an unobserved cause.

8.3.1 Autocorrelation

One type of pattern that has been investigated in previous literature is autocorrelation, when a previous event is statistically correlated with a future event. We will first discuss two types of irrational beliefs about
p. 162 autocorrelation that ↳ people have been shown to endorse and later discuss how beliefs about the underlying unobserved causal mechanisms moderate these fallacies.

One famous example of irrational belief in ‘negative autocorrelation’ is called the gambler's fallacy. Specifically, gamblers often believe that if they lost on the previous gamble (e.g. roulette bet), they are ‘due for a win’ on the next gamble. That is, people sometimes believe that the previous event negatively predicts the next event. A similar phenomenon occurs with other processes that are expected to be random. For example, people think that having six boys in a row (BBBBBB) is less likely than a specific intermixed sequence (e.g. GBBGBG) even though both are equally likely. As famously stated in a Dear Abby column, a woman who just had her eighth girl in a row claimed that ‘this one was supposed to have been a boy’ as if the previous births negatively influenced the chances of the future birth (DEAR ABBY column, reprinted in Hastie & Dawes, 2001, p. 159).

Consider another fallacy termed ‘hot-hand.’ This fallacy is named after the belief that basketball players go through hot streaks of many baskets in a row and cold streaks of many misses in a row. In fact, statistical analyses have not found even a single basketball player whose streaks deviate from chance; however, hot-hand has been found in other sports (Adams, 1995; Dorsey-Palmateer & Smith, 2004; Gilden & Wilson, 1995; 1996; Smith, 2003).

What mediates whether people believe in hot-hand or gambler's fallacy? Restated, for a given random series of events, why do people sometimes infer streaks (positive autocorrelation) and other times infer alternation (negative autocorrelation)? Burns and Corpus (2004) proposed that inferring positive vs. negative autocorrelation depends upon the causal mechanism that people believe to have generated the data. Specifically, people believe that some causal mechanisms have ‘momentum’ and cause streaks, whereas mechanisms that are ‘random’ do not produce streaks. Burns and Corpus presented participants

with scenarios intended to imply either a random mechanism (e.g. roulette wheel) or a mechanism with momentum (e.g. basketball-shooting under competition). Participants were told that in a series of 100 events, the frequencies of the two outcomes had been equal, but that the event ended with a streak of one outcome. In the scenarios that participants believed to be produced by a random mechanism, participants were more likely to predict that the next event would break the streak (i.e. negative autocorrelation), whereas in the conditions that participants believed to have been produced by a non-random mechanism with momentum, participants were more likely to believe that the next event would continue the streak.

Ayton and Fischer (2004) conducted an experiment demonstrating the reverse effect; participants observed a binary sequence of events and were then asked to choose whether they thought the sequence was produced by either a mechanism meant to imply randomness (e.g. roulette wheel) or by a mechanism meant to imply
p. 163 momentum (e.g. basketball shooting). After observing ↳ positive/negative autocorrelation within the sequence of events, people tended to infer non-random/random mechanisms, respectively. In sum, these studies show that people infer different unobserved mechanisms based on different observed patterns.

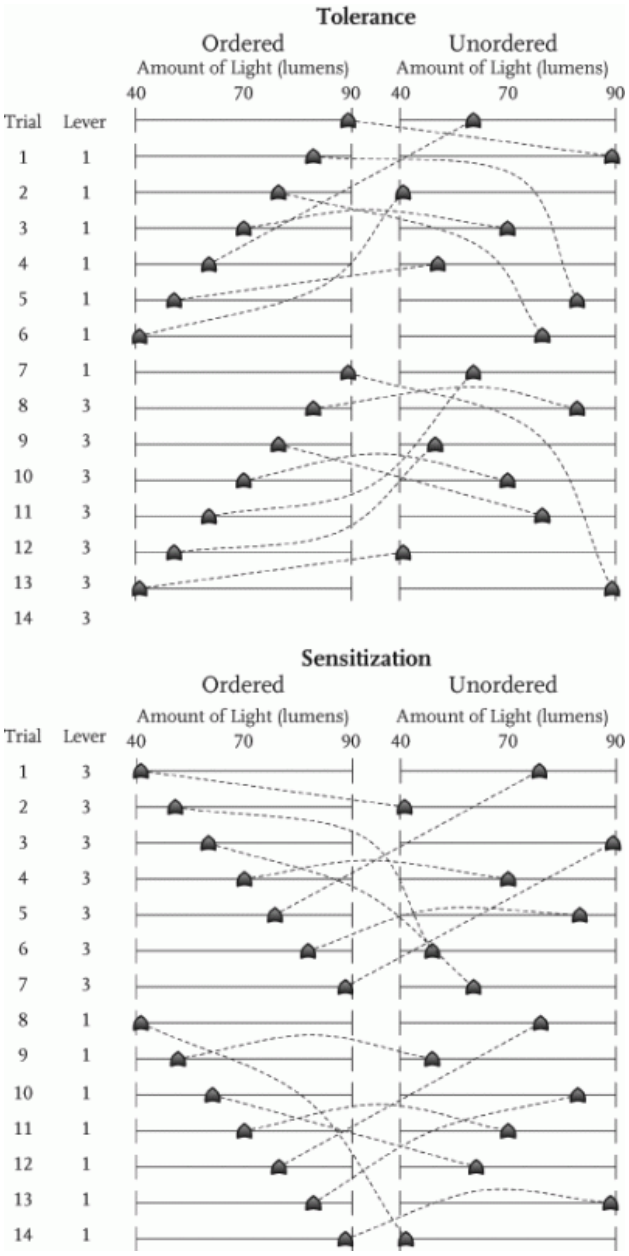
These studies have compared two generating procedures, random and streaky mechanisms. One interesting possibility is that people may also infer a mechanism that frequently alternates. For example, it seems likely that most people do not go to the same restaurant twice in a row—after going to a restaurant once, one would probably switch to a different restaurant for diversity. In a series of data that is strongly alternating, it seems likely that people would infer a third type of mechanism that produces alternating sequences.

8.3.2 Tolerance and sensitization

A recent study by Rottman and Ahn (2009) demonstrates that people infer a causal mechanism given other kinds of patterns: tolerance and sensitization. An example of a *tolerance* scenario is tolerance to coffee. The first time a person drinks one cup of coffee he/she may feel very awake. However, after repeatedly drinking one cup of coffee, he/she becomes tolerant and one cup of coffee has little effect. The person may then drink two cups of coffee and initially feel very alert, but after repeatedly drinking two cups of coffee, again becomes tolerant. In sum, tolerance involves a *decreasing* effect over time when the cause is held constant. *Sensitization* is essentially the opposite of tolerance; sensitization involves an *increasing* effect over time when the cause is held constant. For example, many antidepressants require repeated exposure for full effectiveness. Two pills of antidepressant may initially have no effect, but after repeated exposure, two pills may be sufficient to make a person very happy. If the person cuts down to one pill of antidepressant, the decrease may initially result in a decrease in happiness, but if the person becomes sensitized to the reduced amount of antidepressant, over time, one pill may become sufficient.

To determine whether people are sensitive to these tolerance/sensitization schemata, Rottman and Ahn showed participants scenarios in which machines were tested 14 times in a row for their emissions (e.g. noise, light, heat, or smell). The input to the machines was a lever that could be set to three positions, analogous to the number of cups of coffee or number of pills of antidepressant for the scenarios described above. In one set of 'ordered' conditions, the emissions increased (sensitization) or decreased (tolerance) over repeated use. In another set of 'unordered' conditions, there was no temporal pattern so the data looked random. Figure 8.2 depicts the ordered and unordered versions of both the tolerance and sensitization conditions.

Fig. 8.2



Tolerance and sensitization, ordered and unordered conditions.
Note: Dashed lines (not shown to participants) illustrate how the same data were reordered from the ordered to unordered conditions.

After observing these trials in each condition, participants were asked to 'rate how confident you are that
p. 164 the lever has the capacity to affect' the ↳

p. 165 ↳
emission (e.g. noise, light, heat, or smell). Participants gave considerably higher ratings in the ordered tolerance and sensitization conditions than in the unordered conditions. What is particularly interesting about the results is that people in both the ordered and unordered conditions saw identical data in terms of the simple correlation between the lever and the emission (i.e. not considering the temporal dimension). This is the type of information that has been used as the basis of causal inferences by most traditional causal induction models, and thus, they would have predicted no difference between the two conditions.

Furthermore, the overall correlation was zero. Despite this, people in the ordered condition were moderately confident that the lever had causal efficacy.

One way to explain this finding appeals to unobserved causes. For example, in the ordered-tolerance condition, participants likely inferred a process that occurs within each individual machine such that a latent inhibitory variable increases over time. This is why the emissions decrease over time. In the ordered-sensitization condition, participants likely inferred an unobserved inhibitory cause that decreases over time, explaining why the emissions increase. It seems unlikely that people would think that time itself directly influences the emissions, however, over time, the machine may become 'worn in' and produce less emission. In this case, the variable responsible for 'wearing in' would be the unobserved variable that inhibits the emissions and is correlated with time.

In this account, people would use the temporal pattern to infer an unobserved cause, and the combination of the lever and this unobserved cause completely explains the emissions. This explains why participants judged that the lever influences the emissions. After all, in the ordered conditions, the emission is statistically dependent upon the lever once time or the unobserved cause is taken into account. However, when there is no temporal pattern as in the unordered condition, there is no reason to infer an unobserved cause that changes with time. Consequently people have no way to make sense of the influence of the lever and judge it not causally efficacious.

It is not difficult to find real-world examples of this reasoning. Caffeine is an adenosine antagonist; caffeine inhibits sleep through blocking adenosine, which promotes sleep. However, over repeated caffeine exposure, the number of adenosine receptors increases, making caffeine less effective at blocking them. The number of adenosine receptors is thus an unobserved cause that changes over time within an individual person.

The above situations depict repeated treatments on the same machine. To further test whether people understand the tolerance/sensitization scenarios, we created another set of scenarios in which the increasing/decreasing patterns occur to many different machines. Going back to the coffee scenario, one person's coffee drinking can influence the effectiveness of coffee for that same person at a later time.
p. 166 However, one person's coffee drinking should not ↳ influence the effectiveness of coffee for a different person at a later time—tolerance to coffee must happen within one entity. If people only apply the tolerance/sensitization schemata for one-entity scenarios, then they should give higher causal efficacy ratings for the lever in one-entity scenarios compared to many-entity scenarios that depict the exact same input/output data patterns. In a second experiment designed to test this one-entity vs. many- entity distinction, we found that people were more confident in the causal efficacy of the lever in the one-entity than many-entities conditions both for sensitization and tolerance.

This experiment further clarifies the inferences about the unobserved variable. In this experiment, the data patterns in the one-entity and many-entity conditions were identical for the lever, emissions, and temporal order. However, it is only in the one-entity condition that one can plausibly infer a latent process; an unobserved cause within each individual machine changes and affects the emissions even though the observed cause's strength remains constant. As previously explained, if people infer an unobserved cause in the one-entity scenario, the combination of this unobserved variable and the lever completely accounts for the pattern of emissions, which explains why people rated the lever to be efficacious. However, it would be too bizarre to infer a latent process occurring within each individual getting transferred to the person who happens to drink coffee next or the machine that happens to be tested next. If people do not infer an unobserved cause in the many-entities scenario, the pattern of data between the lever and emission does not make sense; after all, as is also true in the one-entity scenario, there is no simple correlation between the lever and emission. This explains why participants gave lower ratings for the causal efficacy of the lever in the many-entities condition. In sum, this study suggests that inferences about an unobserved cause that

changes over time within one entity influenced inferences about the relationship between the observed cause and effect.

8.3.3 Developmental origins of beliefs about order

The previous studies have focused on cases when adults perceive a pattern and infer an unobserved cause to explain the pattern. Some previous studies have plotted the development of children's beliefs and inferences about causal mechanisms. Friedman (2001) found that four-year-old children believe that it is plausible for animate agents to create an ordered pattern from randomness but less plausible for non-animate causes (e.g. the wind) to do the same. That is, even without seeing a particular cause occur, four-year-olds infer that one type of unobserved cause is more plausible than another.

Newman, Keil, Kuhlmeier, and Wynn (2010) found similar results among even 12-month-old children. They created scenarios in which the infants initially saw either an ordered or unordered pile of blocks. Then an
p. 167 opaque barrier occluded the blocks and either a rolling ball or an animate ↳ agent (a self-propelled circular face) moved behind the occluder, presumably coming in contact with the blocks. Finally, the occluder was removed displaying either ordered and unordered blocks. The infants were more surprised (looked longer) when the ball appeared to create order from disorder than disorder from order, but they looked equally long at the two conditions for the animate agent. In sum, from a fairly early age, children understand that only animate mechanisms can create ordered patterns and infer an unobserved agent to explain an observed pattern.

8.3.4 Other types of patterns and discussion

The tolerance/sensitization experiments described above made us aware that there might be other types of patterns that people may use to infer unobserved causes. When making the unordered tolerance/sensitization conditions (see Figure 8.2), we tried to make the temporal patterns look as random as possible. However, despite our best efforts, in informal discussions after the experiment we discovered that some participants still saw patterns in the data. (See Hastie & Dawes, 2001, p. 355, for an example of the many possible patterns one might infer from a series of six sequential coin flips.) Participants saw increasing or decreasing patterns within subsets of data and interpreted them as meaningful trends (e.g. an increasing pattern in Trials 10–14 in the tolerance-unordered condition in Figure 8.2 despite Trials 9 and 13). Thus, tolerance/sensitization may potentially be triggered for noisier data than what we presented to participants or subsets of data. Some participants also saw alternating patterns (e.g. Trials 1–5 in the tolerance-unordered condition in Figure 8.2) of the form we proposed at the end of Section 8.3.1. Another type of pattern people would likely infer in other situations is a periodic or sinusoidal pattern. A sinusoidal pattern is similar to positive autocorrelation in that the previous trial predicts the next trial, but different in that the period of repetition may be constant which is not necessarily the case for autocorrelation. In all of these scenarios, we believe that people would likely attribute an observed pattern to an unobserved cause, which could further influence their judgments about observed causes. (But see the last section of this chapter for a discussion of boundary conditions.)

There are a number of important future directions of this research. First, it would be useful to determine whether people have a limited set of schemata or patterns they primarily search for when learning new causal relationships. A limited set or taxonomy of plausible schemata could reduce the complexity of causal learning given that there are infinite numbers of possible patterns caused by unobserved variables. Exploring the diversity of causal schemata may help us better understand the limits of causal learning as well as how people make generalizations from schemata they know and learn new schemata.

There might also be important individual differences in inferring unobserved causes due to pattern
p. 168 detection. Certain people, for example, paranoid ↳ people, may have a higher likelihood of seeing a pattern where none exists and attributing the pattern to an unobserved cause. Prior experience with certain types of mechanisms or schemata may also make a person more likely to infer a particular type of mechanism.

Finally, though some of these phenomena (e.g. hot-hand fallacy) have been studied extensively, they are not usually considered the domain of causal reasoning, but rather decision making. It would be useful to integrate research about systematic patterns with more traditional causal learning paradigms that have focused on single events (i.e. the previous section). After all, when observing a new set of data, people are sensitive to both single events and patterns of events, and a general model of causal learning should incorporate both.

8.4 Beliefs in stability of causal relations

The previous two sections have suggested that people infer an unobserved cause to understand unexplained events and systematic patterns of events. In this section, we suggest that people also infer an unobserved cause if they notice that the relationship between an observed cause and effect changes. For example if you know that a medicine has a particular side effect for most people, but find a group of people who do not develop the side effect, it would make sense to infer an unobserved cause to explain the difference (e.g. the group has an unusual gene). Restated, it seems likely that people will infer an unobserved cause when a causal relationship is not stable. We will discuss two types of stability of causal relationships: stability across different samples and stability over time.

8.4.1 Causal power–stability across samples

Cheng (Power PC; Cheng, 1997) proposed that when people judge whether X causes Y, people intuitively estimate *causal power*, the 'probability with which [X] influences [Y]' (Buehner, Cheng, & Clifford, 2003). Consider the following scenario: you are testing the side-effects of a new drug and discover that when given to 100 people without headaches, 50 of these people develop a headache. Suppose you gave the drug to 100 people, 50 of whom already have a headache. How many out of these 100 would have a headache after taking the drug? According to Power PC theory 75 people would have a headache. In the first situation, the medicine caused 50% of the people to get a headache. In the second scenario, 50 people already have a headache, and the medicine will cause 50% of the remaining people to get a headache. The base rate percent of people who already have a headache may vary from situation to situation, but Cheng argues that the percent of people who do not already have a headache and will get a headache should be constant across scenarios.

p. 169 Before moving on, it is useful to understand why Cheng makes this argument, though some readers may prefer to skip ahead to the results of her experiment. One easy way to calculate causal strength is simply to subtract the probability of an effect (E) occurring in the presence of an observed cause (O) minus the probability of the effect occurring in the absence of the cause: $P(E|O) - P(E|\sim O)$ (i.e. ΔP measure mentioned earlier). However, this calculation has the problem that it is influenced by ceiling effects. When the base rate of the effect occurring without the cause is greater than zero, the causal strength cannot be the maximum 1 even if the effect always occurs in the presence of the cause. To get around this problem, Cheng uses a number of assumptions to calculate causal power.[3] First, she assumes that an effect can occur for two reasons: if the observed cause produces the effect or if an unobserved cause (U) produces the effect (or they can both produce it simultaneously). This assumption is very useful because it implies that if E occurs in the absence of O then U must be responsible. Thus the probability of the effect occurring in the absence of the cause, $P(E|\sim O)$ is an estimate of the frequency that the unobserved was responsible for the effect. Second,

she assumes that the observed cause occurs independently from unobserved causes, $P(O|U) = P(O|{\sim}U)$, and that O and U influence E independently. These assumptions are also very useful because we now assume that $P(E|{\sim}O)$ is an estimate of how frequently the unobserved cause produces the effect in general, even when O is present. To determine causal power, ΔP is divided by 1 - $P(E|{\sim}O)$ which effectively normalizes it on the probability that the unobserved cause produced the effect, resulting in the increase in probability of the effect due to the observed cause regardless of unobserved causes.

$$\text{Causal power} = \frac{P(E|O) - P(E|{\sim}O)}{-P(E|{\sim}O)}.$$

The causal power of a particular cause/effect relationship is thus supposed to be the same in samples regardless of unobserved causes (the base rate of E). If different causal powers are observed in different samples, one likely explanation is that the assumptions about the unobserved causes are violated, and that the apparent relationship between the observed cause and effect is partially due to unobserved causes.

Liljeholm and Cheng (2007) tested whether people believe that the causal power of a specific cause is stable across situations. They created two conditions, each of which had three scenarios like the headache scenario. In one condition, the three scenarios had the same causal power but different base rates of headache. In a second condition, the base rate of headache was the same (zero people initially had a
p. 170 headache) but the causal power was different ↳ across the three scenarios. After observing the data for the three scenarios in each condition, participants answered whether they thought the medicine interacts with some unobserved factor across the experiments or whether the medicine has the same influence across the three scenarios. Whereas only one third of participants thought the medicine interacted with an unobserved factor in the causal power constant condition, 86% thought that the medicine interacted with an unobserved factor in the condition in which causal power varied.

In sum, causal power seems to be one way in which people expect causal relations to be stable across scenarios. When it is not stable, people infer an unobserved cause that interacts with the observed cause and is responsible for the discrepant causal power estimates. The type of stability of causal relations discussed here relates to stability across different contexts that are distinguished for learners. That is, Liljeholm and Cheng (2007) presented participants with three scenarios each framed as an individual study with different hypothetical patients. In the next studies, participants learned about changing causal relations on their own.

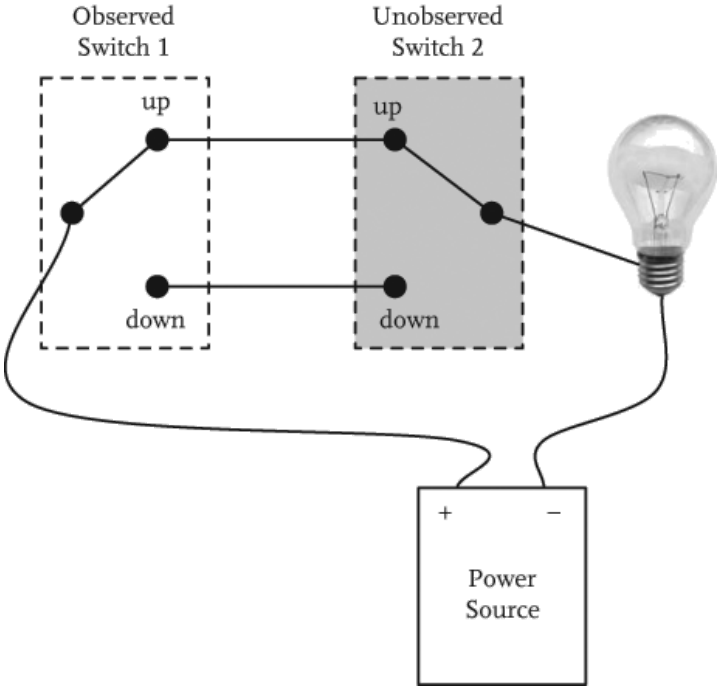
8.4.2 Grouping effects–stability over time

One of the challenging aspects of causal learning is that there are infinitely many possible interacting factors and these interacting factors change over time and context. Sometimes we have a priori beliefs about possible interacting factors; however, often we do not know about interacting factors, or whether interacting factors have changed. In this case, we may learn about interacting factors by noticing a difference in a causal relationship over time. When we observe a change in a causal relationship over time, we will likely conclude that an unobserved interacting factor changed. One critical assumption for making such an inference is that unobserved factors are stable for long enough periods of time that we can notice the difference.

Consider the following double-switch scenario that was briefly discussed at the beginning of this chapter listed in Table 8.1 as *biconditional*. Some lights are connected to two switches (e.g. often at opposite ends of a hallway). There are two important characterizations of this scenario. First, whenever one switch is flipped (assuming that the other light switch is not flipped at exactly the same time), the state of the light will change. Second, neither of the two switches has an ‘on’ or ‘off’ position—there is not necessarily any

correlation between the position of a given switch and the state of the light. Figure 8.3 provides a wiring diagram of a double light switch. The light will be on whenever the switches make a complete circuit (if both switches are up or if both switches are down).

Fig. 8.3



Wiring diagram for double light switch scenario.

Suppose you enter a room for the first time and discover that when you flip a switch up, a light goes on, and when you flip it down, the light goes off (grey cells, Steps 1–4 in Table 8.3). If you assume that other
p. 171 potential causes of the ↳

p. 172 ↳
light are fairly stable (and do not happen to change at the same moment you flipped your switch), you would infer that the switch influences the light. Later (Steps 4–5), the light turns off without anyone touching the switch (perhaps your daughter flipped the other switch unknown to you; U in Table 8.3). Afterwards, when the switch is down, the light is on, and off when up (Steps 5–8). From this scenario, you might be very confident that your switch influences the light; there were two long periods when the status of the switch correlated with the status of the light. Additionally, because the light mysteriously turned off, you might infer an unobserved factor (I in Table 8.3) that interacts with your switch, explaining the overall zero contingency between the switch and light.

Table 8.3 Double light switch, grouped scenario.*

Steps	Switch	Light	U	I	Inferences:
1	0	0	1	1	Switch influenced light.
2	1	1	1	1	
3	0	0	1	1	
4	1	1	1	1	Unobserved factor changed.
5	1	0	0	0	
6	0	1	0	0	Switch influenced light.
7	1	0	0	0	
8	0	1	0	0	

* *Note*: For Tables 8.3 and 8.4, 'U' represents an unobserved interacting factor and 'I' represents a factor learners are likely to infer. 0 represents down for the switch and off for the light, and 1 represents up for the switch and on for the light.

However, inferring the observed switch to be efficacious depends upon the stability of the unobserved cause. For example, consider the same data from Table 8.3, rearranged as in Table 8.4. Initially, the switch is down and the light is off (Step 1). In Step 2 the switch is flipped up, but the light still stays off. In order to believe that the switch is causally efficacious, one must infer that at the moment the switch was flipped, an unobserved factor coincidentally changed and counteracted the effect of the observed switch, as specified under column 'U' (unobserved interacting factor). Then, in Step 3, the light turns on without flipping the switch, and so on. Thus, for the situation shown in Table 8.4, it would be extremely difficult to infer the switch to be causally efficacious: The switch cannot be the sole cause of the light because there is zero contingency with the light. Furthermore, it would be difficult to infer it as part of an interaction because doing so would require inferring an unobserved factor operating as specified under column 'U,' which is counterintuitive; the unobserved interacting factor is highly unstable and exceedingly complicated to track. Instead, the simplest account (intuitively) would be to infer an unobserved factor that is *entirely* responsible for turning the light on and off. Such a factor would be perfectly correlated with the light, as specified under column 'I' (inferred factor). If a learner inferred 'I', he/she would likely infer that the switch is not causally responsible for the light at all.

Table 8.4 Double light switch, ungrouped scenario.

Steps	Switch	Light	U	I	Inferences:
1	0	0	1	0	Switch did not influence light.
2	1	0	0	0	Unobserved factor changed.
3	1	1	1	1	Switch did not influence light.
4	0	1	0	1	Unobserved factor changed.
5	0	0	1	0	Switch did not influence light.
6	1	0	0	0	Unobserved factor changed.
7	1	1	1	1	Switch did not influence light.
8	0	1	0	1	

These two examples were meant to demonstrate that if an unobserved cause is relatively stable for periods
p. 173 of time with a few salient different periods as in ↳ Table 8.3, a learner is likely to infer that an interaction is taking place with an unobserved cause. However, if the scenario is very unstable, as in Table 8.4, then the learner is less likely to infer an interaction with an unobserved cause. Instead, they would likely infer that the unobserved cause is not responsible for the effect at all.

To investigate these inferences, we gave participants a cover-story about machines that produce blocks of various shapes (e.g. square or triangle), and asked participants to determine if the position of a lever on the machines affects the shape of the blocks. Participants then observed videos of 20 trials in a continuous

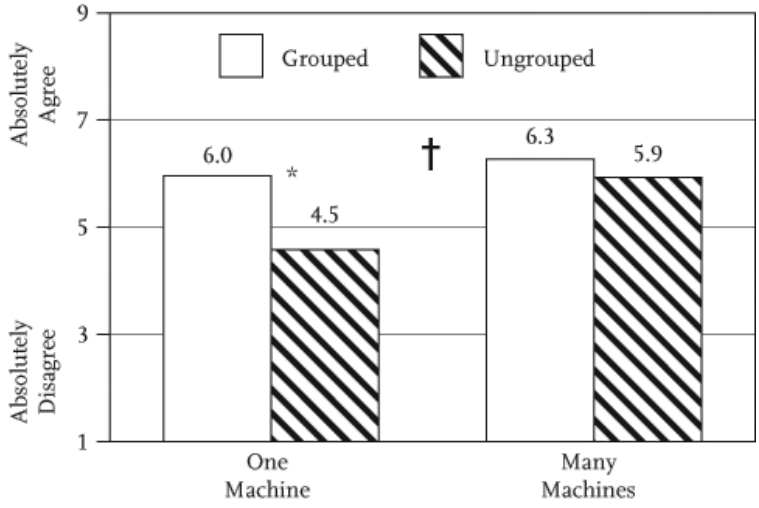
temporal sequence; from trial to trial, the lever sometimes switched between the left and right position and the shape of block (e.g. square or triangle) sometimes changed. In all conditions, the lever was statistically uncorrelated with the shape of the blocks.

Rottman and Ahn (in prep. see Rottman & Ahn 2009, for partial results) manipulated two aspects of the scenarios. First, we manipulated the grouping of the trials similar to that shown in Tables 8.3 and 8.4. In the 'grouped' condition, there were relatively stable periods of time when one shape was associated with one position of the lever, and other periods when the association flipped. In the 'ungrouped' condition, these two different associations were more intermixed so that there were no discernable stable periods. If grouping allows people to infer an unobserved cause that is stable for periods of time and then switches, people should infer an interacting unobserved cause more in the grouped than ungrouped condition.

Another manipulation of the study was whether the scenario involved only one machine changing over time or different machines. In the one- machine condition, all 20 trials occurred with one machine. That is, in the one-machine conditions, the lever on the machine was sometimes flipped back and forth between left and right, and the shape of the block produced by the machine sometimes changed over 20 trials. In the many-machines conditions, 20 different machines were observed once each: the lever of each machine was set either to the left or the right, and the machine produced either a square or triangle. Even though the many-machine conditions were identical to the one-machine conditions in every other way, we reasoned that participants would not make different inferences about the unobserved cause between the grouped vs. ungrouped conditions. Because each machine is different, we reasoned that participants would not make use of the temporal grouping information to infer a stable unobserved cause; after all, participants had no reason why they were presented with the machines in the particular order. If the temporal stability information was not deemed important and people did not use it to infer stability, there should not be any difference between the grouped and ungrouped conditions for many- machines. Such a finding would suggest a caveat to the assumption of stability: people only distinguish between stable and unstable
p. 174 scenarios for ↳ inferring interactions with unobserved causes when time is a meaningful variable.

After observing each scenario, participants rated their agreement with whether 'A combination of the lever and some other factor influenced the shape of the blocks' from 1 ('Absolutely Disagree') to 9 ('Absolutely Agree') (see Figure 8.4). As expected, in the one-machine condition, participants inferred an interaction with an unobserved cause more in the grouped than ungrouped condition. However, in the many-machine condition, when time was not a meaningful factor, there was no difference between the grouped and ungrouped conditions in participants' inferences about an interaction with an unobserved cause.[4] In other words, only when time is a meaningful variable (i.e. the one-machine condition), do people use temporal stability to infer an interaction with an unobserved cause. When time is not meaningful (i.e. the many-machine condition), there is no difference between grouped and ungrouped conditions.

Fig. 8.4



Combination of lever and unobserved factor influenced shape.
Note: * paired t-test, $p < 0.05$,† interaction in repeated-measures ANOVA, $p < 0.05$.

Inferring that there is an unobserved interacting factor has further implications for peoples' views of the observed cause. Specifically, ↳ Rottman & Ahn (2009; Experiment 1) demonstrated that the more grouped the scenario, the higher the causal strength ratings that participants gave it. Across these two experiments, when a scenario is grouped, people are able to infer the interacting unobserved cause and still believe that the observed cause influences the effect even though there is no correlation between the two. However, when the scenario is ungrouped, people are less likely to infer an unobserved factor and more likely to infer that the observed cause is not related to the effect (after all, there is zero correlation).

p. 175

One of the important implications of this study is that people spontaneously distinguish between scenarios with stable unobserved causes and scenarios with unstable unobserved causes, even though both can appear to have zero correlation between the cause and effect. Consider the study briefly discussed in the introduction that investigated the role of pollution (observed cause) on daily temperature range (effect; Gong, Guo, & Ho, 2006). The researchers found that pollution *decreases* daily temperature range during the winter, but pollution *increases* daily temperature range during the summer. Summarized, the season flips the direction of the influence of pollution on diurnal temperature range. This example makes an important point: the researchers did not know about the interaction with season a priori. At some point, they must have noticed that the relationship between pollution and weather is flipped depending on the season, and if they had overlooked this important factor, the relationship between pollution and diurnal temperature range would have been obscured and might appear to not exist. By observing a stream of data and noticing periods of stability, the researchers could uncover that an unpredicted variable (season) plays an important interacting role.

In summary, these studies suggest that people believe causal relationships to be fairly stable across contexts, and if they notice a difference across samples or times, they posit an unobserved factor to explain the difference. Presumably people would also infer systematic patterns to be stable across different contexts, and would likewise posit an unobserved factor if they notice a difference. For example, if a cause exhibited tolerance for one sample and sensitization for another, or positive autocorrelation for one sample and negative autocorrelation for another, people would likely infer an unobserved cause to explain the difference. Additionally, if a cause appeared to switch from a noisy-or to a biconditional functional relationship, people would also infer that an unobserved factor changed. In this way, the inferences we have discussed about single events, patterns of events, and relationships between causes/effects can be viewed in a hierarchy. If an unexplained change occurs anywhere along the hierarchy from the lowest level (single

events), middle level (patterns of events) or the highest level (relationships between causes/effects), people will infer that an unobserved cause is responsible.

p. 176

8.5 Working memory–A requirement for reasoning about unobserved causes

We have now discussed a number of situations when people infer unobserved causes. However, reasoning about unobserved causes is also cognitively challenging. As we have already explained, there are many possible unobserved and unattended causes, and many ways in which those causes can interact with observed causes. In this section, we propose that reasoning about unobserved causes requires considerable working memory capacity. We will now review a particular phenomenon, recency/primacy effects, in which beliefs about unobserved causes play a central role. Then we will demonstrate how working memory mediates this phenomenon.

Suppose you initially observe a set of data, mostly showing positive covariation between two events, followed by data mostly showing negative covariation between the same events. For instance, for the first half of the baseball season, you notice that your favorite baseball team was more likely to win when you were wearing your 'lucky' socks, but for the second half of the baseball season, you notice that your team was more likely to lose when you were wearing your 'lucky' socks. Would you consider your socks to be still lucky? There are many possible strategies a reasonable learner could take to answer this question. One could average across all of the available data, concluding that the socks have nothing to do with winning. Alternatively, one could give more weight on the most recent data, concluding that wearing those socks actually hurts performance. Or one can give more weight on initial data, concluding that wearing those socks improves performance. It is difficult to tell which one of these three is the most rational strategy, and in fact, the experimental results using this paradigm show that people demonstrate all three strategies (Dennis & Ahn 2001; Glautier, 2008; Lòpez, Shanks, Almaraz, & Fernàndez, 1998; Marsh & Ahn, 2006; Shanks, Lòpez, Darby, & Dickinson, 1996).

In a recent study, Luhmann and Ahn (in press) found empirical evidence that the conflicting findings of primacy/recency effects can be explained by learners' beliefs about unobserved causes. Specifically, people who reason more about unobserved causes tend to show primacy, whereas people who reason less about unobserved causes tend to show recency effects in their causal strength judgments. Learners were presented with sequences of covariation information involving medications and potential side effects. Sequences always used the same set of observations, but were constructed to present the majority of positive evidence first, followed by the majority of negative evidence (Positive—Negative) or vice versa (Negative—Positive). Sometimes during the learning sequence, participants were asked to explain why the effect did or did not occur, and at the end of the sequence, learners made causal strength judgments.

p. 177

Some subjects were particularly likely to explain the outcome of a specific trial by appealing to unobserved causes. For example, consider a learner in the Positive–Negative condition who observed the first half of positive evidence, and then observes some contradictory negative evidence (i.e. a trial when the cause occurs but the effect does not). Participants were then prompted to choose one explanation for why this happened: '[the cause] prevented [the effect]', 'it is pure coincidence that' [the effect] did not occur after [the cause]', or 'for some reason, [the cause] failed to cause [the effect].' If the participant appeals to an unobserved cause, he/she would choose the third option, which subtly references alternative causal influences (i.e. 'for some reason'). The unobserved cause could have overridden the observed cause and prevented the effect, allowing participants to continue to believe that the observed cause was generative. In fact, participants who choose this option gave higher causal strength ratings in the Positive–Negative condition. (These participants exhibited a primacy effect because their higher causal strength ratings reflect the initial

positive contingency.) These results suggest that unexpected covariation information elicits reasoning about unobserved, alternative causes in some learners. Such reasoning tends to 'excuse' the new, contradictory information and leave the prior causal beliefs relatively untouched.

However, other learners appealed to unobserved causes less and instead used this same conflicting information to directly modify their causal beliefs. For example, in the Positive—Negative case, upon encountering the negative evidence, a learner could take this evidence at face value and modify the initial belief that the cause generates the effect to conclude that the cause is not related to the effect or even prevented the effect. Such learners subsequently gave lower causal strength estimates (a recency effect).

So far, these results suggest that primacy/recency effects in causal judgments are related to whether people appeal to unobserved causes. However, why do some people appeal to unobserved causes more than others? We hypothesize that one of the reasons is the ease with which learners are able to reason about unobserved causes that produces these different learning strategies. Marsh and Ahn (2006) demonstrated that learners with higher verbal working memory capacity were more likely to show a primacy effect. Thus, we reasoned that working memory may facilitate reasoning about unobserved causes, which we know influences primacy/recency in causal strength judgments. To test this hypothesis, Luhmann and Ahn (in press) created two situations that experimentally manipulated the ease of reasoning about unobserved causes.

First, Luhmann and Ahn (in press) increased the cognitive load during the task, which we predicted would decrease the ease of reasoning about unobserved causes and produce a recency effect. The learners performed the same task explained above while simultaneously performing a difficult secondary task
p. 178 (counting backwards by 3s) This manipulation impaired rea- ↳ soning about unobserved causes. Participants took individual trials at face value and modified existing hypotheses. For example, in the Positive–Negative condition, when faced with the negative evidence, participants simply said that the cause inhibits the effect, presumably because it would be too taxing on working memory to postulate unobserved causes. Furthermore, participants' overall causal strength judgments showed a recency effect.

Second, we made it easier for learners to reason about unobserved causes by simply making the unobserved causes observed. During the second, contradictory half of the event sequence, learners were told that an alternative cause was present. Note that they were not told that the effect occurred because of this alternative cause (that is, there still is an ambiguity as to what was the true cause of the effect). Yet, this manipulation increased participants' reasoning about unobserved causes and they interpreted information that contradicted their prior causal beliefs as "something" going wrong, leaving their beliefs relatively untouched. Furthermore, their overall causal strength judgments showed a primacy effect.

In sum, these studies demonstrate that working memory moderates inferences about unobserved causes for primacy/recency effects. Based on these results, it is plausible that working memory would moderate other inferences about unobserved causes, such as those in the double light switch scenario or perhaps those discussed in the section on Power PC. Given that reasoning about unobserved causes (e.g. potential confounds; see Cheng, 1997) is necessary for normative causal inference, future research on the limits and conditions under which people reason about unobserved causes is particularly important.

8.6 Conclusions

In this chapter, we demonstrated that people make a number of sophisticated inferences about unobserved causes. First, people infer an unobserved cause when a single unexplained event happens: children appeal to magic when they don't have a physical explanation for why one object would move without another object touching it (Phelps & Woolley, 1994), and adults infer that a hidden button was pressed when a light bulb illuminates without an observed button being pressed (Luhmann & Ahn, 2007). Second, people infer an unobserved factor to explain patterns of events that cannot be explained by the observed causes, such as a series of 10 coin flips all landing on heads (Ayton & Fischer, 2004), or a person becoming tolerant to caffeine (Rottman & Ahn, in press). (In both these scenarios, the mechanism that produced the pattern is not observed.) Third, people infer an unobserved cause to explain changes in the relationship between a cause and effect, for example, if a cause sometimes generates the effect and sometimes inhibits the effect (Rottman & Ahn, 2009).

p. 179 Though we have not focused much on causal learning models, these experiments suggest some important implications for the development of future models. Most existing models of causal learning have focused upon observed causes and make fairly simple assumptions about unobserved causes. For example, existing models assume that unobserved causes are always present (Rescorla & Wagner, 1972) or are not confounded with observed causes (e.g. Cheng, 1997; see Luhmann & Ahn, 2007, for a discussion). However, people do not seem to make these assumptions and instead make dynamic inferences about unobserved causes; people do not believe that unobserved causes are constant and make sophisticated inferences about the presence and changes in unobserved causes.

Many of the phenomena discussed in the current chapter involve scenarios that unfold over time (i.e. autocorrelation, tolerance/sensitization, the double light switch scenario, and primacy/recency effects), yet most existing models do not capture time sufficiently. Many influential models have been designed primarily to capture causal phenomena that do not occur over time, and thus they aggregate across all trials (e.g. Cheng, 1997; Jenkins & Ward, 1965; Griffiths & Tenenbaum, 2005). As already discussed, influential animal-learning models that do model learning over time (e.g. Rescorla & Wagner, 1972) often make overly-simplistic assumptions about unobserved causes. Bayesian inference has become a particularly common way to model causal learning, and it proved very useful in BUCKLE. However, as noted by Danks (2007), 'causal Bayes nets do not currently provide good models of continuous time phenomena, though continuous time Bayes nets are the subject of ongoing research (Nodelman 2002, 2003)'. We believe that capturing phenomena that unfold over time should be an important aspect of future models.

We have also demonstrated that when working memory is taxed, people have difficulty reasoning about unobserved causes (Luhmann & Ahn, in press) Yet, despite the additional cognitive challenge of reasoning about unobserved causes, we believe that it occurs as a normal part of causal learning. Many of the previously mentioned studies have demonstrated that people spontaneously reason about unobserved causes. Furthermore, we have described multiple phenomena in which reasoning about unobserved causes influences the inferences people make about observed causes, the more typically studied form of causal learning. For example, when people infer that an unobserved cause flips the relationship between the observed cause and effect (generative vs. preventative), since the interaction explains why there may be zero overall correlation, people infer that the observed cause still influences the effect (Rottman & Ahn, 2009).

One general way to summarize these findings is that people are always on the lookout for unexplained data and consequently infer unobserved causes to make sense of the relationship between observed causes and
p. 180 effects. Any attempt to explain causal learning based on overly simplistic and static ↳ assumptions about unobserved causes will not be able to account for the dynamic interplay between observed and unobserved

causes demonstrated in these studies. Perhaps reasoning about unobserved causes should be viewed as a fundamental feature of 'causal' reasoning.

Acknowledgements

This research was supported by an NSF Graduate Research Fellowship (Rottman) and NIMH Grant MH57737 (Ahn). The authors thank two anonymous reviewers for useful suggestions.

References

Adams, R. M. (1995). Momentum in the performance of professional tournament pocket billiards players. *International Journal of Sport Psychology*, *26*, 580–587.
WorldCat

Arkes, H. R. & Harkness, A. R. (1983). Estimates of contingency between two dichoto- mous variables. *Journal of Experimental Psychology: General*, *112*, 117–135. 10.1037/0096-3445.112.1.117
WorldCat Crossref

Audi, R. (1995). *The Cambridge Dictionary of Philosophy*. Cambridge, MA: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Ayton, P. & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition*, *32*, 1369–78. 10.3758/BF03206327
WorldCat Crossref

Burns, B. D. & Corpus, B. (2004). Randomness and inductions from streaks: 'Gambler's fallacy' versus 'hot hand'. *Psychonomic Bulletin & Review*, *11*, 179–84. 10.3758/BF03206480
WorldCat Crossref

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405. 10.1037/0033-295X.104.2.367
WorldCat Crossref

CNN (1999, May 13). Night-light may lead to nearsightedness. Retrieved June 25, 2009, from http://www.cnn.com/HEALTH/9905/12/children.lights/index.html

Danks, D. (2007). Causal learning from observations and manipulations. In M. C. Lovett & P. Shah (Eds.), *Thinking with Data* pp. 359–388. New York: Lawrence Erlbaum Associates
Google Scholar Google Preview WorldCat COPAC

Dennis, M. J. & Ahn, W. K. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition*, *29*, 152–64. 10.3758/BF03195749
WorldCat Crossref

Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgment of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, *36A*, 29–50.
WorldCat

Dorsey-Palmateer, R. & Smith, G. (2004). Bowlers' hot hands. *The American Statistician*, *58*, 38–45. 10.1198/0003130042809
WorldCat Crossref

Downing, C. J., Steinberg, R. J., & Ross, B. H. (1985). Multicausal inference: Evaluation of evidence in causally complex situations. *Journal of Experimental Psychology: General*, *114*, 239–263. 10.1037/0096-3445.114.2.239
WorldCat Crossref

Einhom, H. J. & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*, 3–19. 10.1037/0033-2909.99.1.3
WorldCat Crossref

Friedman, W. J. (2001). The development of an intuitive understanding of entropy. *Child Development*, *72*, 460–73. 10.1111/1467-8624.00290
WorldCat Crossref

Gilden, D. L. & Wilson, S. G. (1995). On the nature of streaks in signal-detection. *Cognitive Psychology*, *28*, 17–64. 10.1006/cogp.1995.1002
WorldCat Crossref

Gilden, D. L. & Wilson, S. G. (1996). Streaks in skilled performance. *Psychonomic Bulletin & Review*, *2*, 260–265. 10.3758/BF03210967
WorldCat Crossref

Gilovich, T. (1991). *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*. New York: Free Press.
Google Scholar Google Preview WorldCat COPAC

Glautier, S. (2008). Rcency and primacy in causal judgments: Effects of probe question and context switch on latent inhibition and extinction. *Memory & Cognition*, *36*, 1087–93. 10.3758/MC.36.6.1087
WorldCat Crossref

Goldvarg, E. & Johnson-Laird, P. N. (2001). Naive causality: A mental model theory of causal meaning and reasoning. *Cognitive Science*, *25*, 565–610. 10.1207/s15516709cog2504_3
WorldCat Crossref

Gong, D., Guo, D., & Ho, C. (2006). Weekend effect in diurnal temperature range in China: Opposite signals between winter and summer. *Journal of Geophysical Research*, *111* D18113, doi:10. 1029/2006JD007068.
WorldCat

Griffiths, T. L. & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*, 334–384. 10.1016/j.cogpsych.2005.05.004
WorldCat Crossref

Gwiazda, J., Ong, E., Held, R., & Thorn, F. (2000). Vision: Myopia and ambient nighttime lighting, *Nature*, *404*, 144. 10.1038/35004663
WorldCat Crossref

Hagmayer, Y. & Waldmann, M. R. (2007). Inferences about unobserved causes in human contingency learning. *Quarterly Journal of Experimental Psychology (2006)*, *60*, 330–55. 10.1080/17470210601002470
WorldCat Crossref

Hastie, R. & Dawes, R. M. (2001). *Rational Choice in an Uncertain World*. Thousand Oaks, CA: Sage.
Google Scholar Google Preview WorldCat COPAC

Jenkins, H. M. & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied*, *79*, 1–17. 10.1037/h0093874
WorldCat Crossref

Koslowski, B. (1996). *Theory and Evidence: The Development of Scientific reasoning.* Cambridge, MA: MIT Press.
Google Scholar Google Preview WorldCat COPAC

Leising, K. J., Wong, J., Waldmann, M. R., & Blaisdell, A. P. (2008). The special status of actions in causal reasoning in rats. *Journal of Experimental Psychology: General*, *127*, 514–527. 10.1037/0096-3445.137.3.514
WorldCat Crossref

Liljeholm, M. & Cheng, P. W. (2007). When is a cause the 'same'? Coherent generalization across contexts. *Psychological Science*, *18*, 1014–21. 10.1111/j.1467-9280.2007.02017.x
WorldCat Crossref

Lòpez, F. J., Shanks, D. R., Almaraz, A., & Fernàndez, P. (1998). Effects of trial order on contingency judgments: A comparison of associative and probabilistic contrast accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 672–

694. 10.1037/0278-7393.24.3.672
WorldCat Crossref

Luhmann, C. C. & Ahn, W. (in press). Order Effects during Learning: Expectations and Interpretations, *Journal of Experimental Psychology: Learning, Memory, and Cognition*

Luhmann, C. C. & Ahn, W. K. (2005). The meaning and computation of causal power:Comment on Cheng (1997) and Novick and Cheng (2004). *Psychological Review*, *112*, 685–93. 10.1037/0033-295X.112.3.685
WorldCat Crossref

Luhmann, C. C. & Ahn, W. K. (2007). BUCKLE: A model of unobserved cause learning. *Psychological Review*, *114*, 657–77. 10.1037/0033-295X.114.3.657
WorldCat Crossref

Marsh, J. K. & Ahn, W. K. (2006). Order effects in contingency learning: The role of task complexity. *Memory & Cognition*, *34*, 568–76. 10.3758/BF03193580
WorldCat Crossref

p. 182

Newman, G. E., Keil, F. C., Kuhlmeier, V., & Wynn, K. (2010) Sensitivity to design: Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences*.

Nisbett, R. E. & Ross, L. (1980). *Human Inference: Strategies and Short-comings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.
Google Scholar Google Preview WorldCat COPAC

Nodelman, U., Shelton, C. R., & Koller, D. (2002). Continuous time Bayesian networks. In *Proceedings of the 18th international conference on uncertainty in artificial intelligence*, pp. 378–387.
WorldCat

Nodelman, U., Shelton, C. R., & Koller, D. (2003). Learning continuous time Bayesian networks. In *Proceedings of the 19th international conference on uncertainty in artificial intelligence*, pp. 451–458.

Novick, L. R. & Cheng, P. W. (2004). Assessing interactive causal influence. *Psychological Review*, *111*(2), 455–85. 10.1037/0033-295X.111.2.455
WorldCat Crossref

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Phelps K. E. & Woolley, J. D. (1994). The form and function of children's magical beliefs. *Developmental Psychology*, *30*, 385–394. 10.1037/0012-1649.30.3.385
WorldCat Crossref

Quinn, G., Shin, C., Maguire, M. & Stone, R. (1999). Myopia and ambient lighting at night, *Nature*, 399, 113–114. 10.1038/20094
WorldCat Crossref

Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (eds.), *Classical Conditioning II: Current Research and Theory*. New York: Appelton-Century-Crofts.
Google Scholar Google Preview WorldCat COPAC

Rottman, B. M. & Ahn, W. (2009). Causal inference when observed and unobserved causes interact. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*.

Rottman, B. M. & Ahn, W. K. (2009). Causal learning about tolerance and sensitization. *Psychonomic Bulletin and Review*, 16(6),

1043–9. 10.3758/PBR.16.6.1043
WorldCat Crossref

Rottman, B. M. & Ahn, W. (in prep.) Effect of Grouping of Evidence types on Learning about interactions between observed and unobserved causes.

Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, *16*, 995–1001. 10.1111/j.1467-9280.2005.01649.x
WorldCat Crossref

Saxe, R., Tzelnic, T., & Carey, S. (2007). Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. *Developmental Psychology*, *43*, 149–58. 10.1037/0012-1649.43.1.149
WorldCat Crossref

Schustack, M. W. & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, *110*, 101–120. 10.1037/0096-3445.110.1.101
WorldCat Crossref

Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory and Cognition*, *8*, 459-467. 10.3758/BF03211142
WorldCat Crossref

Shanks, D. R., Lòpez, F. J., Darby, R. J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (eds.), Causal Learning: The Psychology of *Learning and Motivation*. pp. 265–311. San Diego, CA: Academic Press.
Google Scholar Google Preview WorldCat COPAC

Smith, G. (2003). Horseshoe pitchers' hot hands. *Psychonomic Bulletin & Review*, *10*, 753–758. 10.3758/BF03196542
WorldCat Crossref

Wegner, D. M. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
Google Scholar Google Preview WorldCat COPAC

p. 183 White, P. A. (2005). The power PC theory and causal powers; Reply to Cheng (1997) and Novick and Cheng (2004). *Psychological Review*, *112*, 675–682. 10.1037/0033-295X.112.3.675
WorldCat Crossref

White, P. A. (2009). Not by contingency: Some arguments about the fundamentals of human causal learning. *Thinking & Reasoning*, *15*, 129–166. 10.1080/13546780902734236
WorldCat Crossref

Woolley, J. D. (1997). Thinking about fantasy: Are children fundamentally different thinkers and believers from adults? *Child Development*, *68*, 991–1011. 10.2307/1132282
WorldCat Crossref

Notes

1 Here we distinguish causal determinism, the idea that every event must have a cause, from deterministic causality, the belief that an effect must be present if its cause is present (Goldvarg & Johnson-Laird, 1994; Koslowski, 1996; Luhmann & Ahn, 2005).

2 To compute this measure of Δ*P*, the quantities *P*(*E*|*U*) and *P*(*E*|~*U*) were derived by applying Bayes rule to the quantities *P*(*U*|*E*) and *P*(*U* |Δ *E*) which were computed using learners' trial-by-trial likelihood judgments. In this way, the resulting Δ *P* measures the extent to which each subject believed the unobserved cause covaried with the effect.

3 See Cheng (1997) for the complete, formal treatment. Note that we use different notation for consistency within this chapter.

4 As seen in Figure 8.4, participants were more likely to think that there was an interaction with an unobserved cause in the many-machines condition than in the one-machine condition. The reason for this finding is likely because they thought the different machines interact differently with the switch — the different machines is the second interacting factor.