Journal of Experimental Psychology:
Learning, Memory, and Cognition
2015, Vol. 41, No. 2, 395–416

0278-7393/15/$12.00 http://dx.doi.org/10.1037/a0038385

Temporal and Statistical Information in Causal Structure Learning

Teresa McCormack
Queen's University Belfast

Caren Frosch
University of Leicester

Fiona Patrick
University College London

David Lagnado
University College London

Three experiments examined children's and adults' abilities to use statistical and temporal information to distinguish between common cause and causal chain structures. In Experiment 1, participants were provided with conditional probability information and/or temporal information and asked to infer the causal structure of a 3-variable mechanical system that operated probabilistically. Participants of all ages preferentially relied on the temporal pattern of events in their inferences, even if this conflicted with statistical information. In Experiments 2 and 3, participants observed a series of interventions on the system, which in these experiments operated deterministically. In Experiment 2, participants found it easier to use temporal pattern information than statistical information provided as a result of interventions. In Experiment 3, in which no temporal pattern information was provided, children from 6- to 7-years-old, but not younger children, were able to use intervention information to make causal chain judgments, although they had difficulty when the structure was a common cause. The findings suggest that participants, and children in particular, may find it more difficult to use statistical information than temporal pattern information because of its demands on information processing resources. However, there may also be an inherent preference for temporal information.

Keywords: causal learning, time, causal Bayes nets

Historically, most research on causal learning has focused on the cues used to infer whether a particular object or event type is causally efficacious (e.g., Cheng, 1997; Shanks, 2010; Shultz, 1982). However, in real-world reasoning we often have to infer not just the presence or absence of a single causal property, but the structure of causal relations between multiple variables (Waldmann, Hagmayer, & Blaisdell, 2006). Using an example from Gopnik and Schulz (2007), we might observe that every time we go to a party, we have a drink, and have insomnia. What is the causal structure of the relations between the events of going to the party, drinking, and having insomnia? One possibility is that going to the party is a common cause of drinking and insomnia (going to the party causes you to drink, and independently, causes you to have insomnia); an alternative is that going to the party causes drinking which in turn causes insomnia—a causal chain. These alternative causal models, depicted in Figure 1, are just some of those that are possible when three events A, B, and C co-occur, and observing the reliable co-occurrence of these events does not in itself allow us to decide which one is correct. As Gopnik and Schulz (2007) point out, one important way to discriminate between models is to manipulate one of the variables in the system. In this example, one could intervene on drinking (Event B), and deliberately fix its value to drinking or nondrinking, irrespective of party attendance (Event A). This intervention would allow us to discriminate between the common cause and causal chain models shown in Figure 1: If intervening on whether or not B (drinking) occurs has no impact on whether or not C (insomnia) occurs, then we can rule out the causal chain model ABC.

More recently, a variety of studies have examined people's ability to discriminate between different causal structures, and the cues that are used to make such discriminations (e.g., Hagmayer, Sloman, Lagnado, & Waldmann, 2007; Kushnir, Gopnik, Lucas, & Schulz, 2010; Lagnado & Sloman, 2004; Rottman & Keil, 2012; Sobel & Kushnir, 2006; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003; Taylor & Ahn, 2012; White, 2006). Many of these studies have focused on two types of information: statistical information, usually provided as a result of intervening on variables in the system, and temporal order information, which results from observing the temporal pattern of event occurrence.

Statistical Information

The ability to use statistical information in order to derive causal structure has been taken as evidence for the causal Bayes net approach to causal inference, which emphasizes the use of such information to distinguish between competing hypotheses about

This article was published Online First December 22, 2014.

Teresa McCormack, School of Psychology, Queen's University Belfast; Caren Frosch, School of Psychology, University of Leicester; Fiona Patrick, and David Lagnado, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London.

This research was supported by Grant RES-062-23-1799 from the Economic and Social Science Research Council, United Kingdom

Correspondence concerning this article should be addressed to Teresa McCormack, School of Psychology, Queen's University Belfast, Belfast, Northern Ireland, BT7 1NN. E-mail: t.mccormack@qub.ac.uk

Figure 1. Common cause and causal chain structures.

causal structure (Gopnik et al., 2004; Hagmayer et al., 2007; examples of more recent Bayesian accounts of causal inference are described in Griffiths & Tenenbaum, 2009; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008). Within the causal Bayes net approach, causal learning is modeled in terms of the formation and updating of causal graphs. When a probabilistic system operates, information such as the conditional and unconditional dependencies between sets of variables will be available from simply observing the system (i.e., in the absence of intervention), and this information will typically provide evidence for some causal structures rather than others.

Based on the causal Bayes net framework, there are two main computational approaches to structure learning from statistical data: *Bayesian* methods and *constrained-based* methods. On the Bayesian approach (Griffiths & Tenenbaum, 2009; Heckerman, 1998), the central task is to identify how likely each possible structure is given the observed data (i.e., the posterior probability distribution over structures). Bayes' rule is used to update the prior probability distribution (over structures) based on the likelihood of the observed data given each possible structure (see below for more details). When the model parameters[1] are unknown, the Bayesian approach integrates over possible parameter values, to yield a posterior distribution over structures without committing to a specific set of parameter values (Griffiths & Tenenbaum, 2009). In other words, the Bayesian approach assumes prior beliefs about the probabilities of the different structures (possibly that all structures are equally likely), and in the light of the evidence (the observed data), updates these beliefs based on how likely the data is judged to be, given each possible structure.

In contrast, constraint-based methods (Spirtes, Glymour, & Scheines, 1993; see also Gopnik et al., 2004) first identify all dependencies or independencies between variables, using statistical tests on the observed data. This includes testing for patterns of conditional independence: For example, whether or not variable A is independent of variable C conditional on variable B. Given these tests, constraints are placed on the possible structures, as different classes of structure give rise to different patterns of dependency. For example, each of the three causal structures used in the current experiments yield a different set of conditional independence relations (see below). Note that based on observational data alone this method only discriminates *Markov-equivalent* causal structures. Thus, a chain Model A→B→C and a common cause Model A←B→C share the same conditional independencies (namely, A is independent of C conditional on B) and thus are indistinguishable via constraint-based methods alone.

Although initially developed as machine learning methods, both of these approaches have been suggested as computational models of human learning. For example, Gopnik et al. (2004) argue for constraint-based methods, whereas Griffiths and Tenenbaum (2009) argue for Bayesian methods. In general, however, both approaches are computationally demanding, especially for systems with more than three variables. Thus, Bayesian approaches require updating across a large hypothesis space of possible models, and constraint-based approaches require large datasets for identifying conditional and unconditional dependencies. This renders them less plausible as psychological models, and several heuristic approaches to structure learning have been proposed (Fernbach & Sloman, 2011; Mayrhofer & Waldmann, 2011). Moreover, experimental studies have found that when people rely on observational data alone, performance is relatively poor compared to a Bayesian or constraint-based model, even for simple three-variable systems (Lagnado & Sloman, 2004; Steyvers et al., 2003).

Interventional Learning

A key advantage of the causal Bayes net approach is that it can model updating that occurs not just on the basis of observing a system, but also as a result of making interventions. In this approach, the notion of an intervention is formally defined as an external cause that selectively fixes the value of the intervened-on variable, and the effects of such interventions can be modeled using the "do" calculus (Hagmayer et al., 2007; Pearl, 2000; Sloman & Lagnado, 2005). The probabilistic information that is provided as a result of such interventions may provide evidence that the system has a specific causal structure. To return to our example of parties (A), drinking (B), and insomnia (C); if C does not occur if B is prevented from occurring, but B still occurs even if C is prevented, this suggests that the system is an ABC causal chain. Indeed causal learning is improved when participants are provided with probabilistic information that results from interventions on the causal system (Lagnado & Sloman, 2004, 2006; Sobel & Kushnir, 2006; Steyvers et al., 2003).

Temporal Information

Another important type of information in causal structure learning is temporal information (Burns & McCormack, 2009; Frosch, McCormack, Lagnado, & Burns, 2012; Lagnado & Sloman, 2004, 2006; White, 2006). A series of studies have shown that adults use simple temporal heuristics in order to distinguish between causal structures (Lagnado & Sloman, 2004, 2006). Lagnado and Sloman have pointed out that the temporal order in which events occur is often a basic but very useful cue when deciding the causal structure of events. They provide evidence that adults will systematically use a simple temporal heuristic that assumes that correlated changes that occur after one event are likely to be caused by it (see Rottman & Keil, 2012, for a related suggestion, and Rottman, Kominsky, & Keil, 2014, for evidence that even young children use such cues). Thus, if events occur in the temporal order A, followed by B, followed by C, applying such a heuristic iteratively will lead to a judgment that the structure is causal chain ABC, such as that depicted in Figure 1. Burns and McCormack (2009) also demonstrated that if adults observe a three-variable system in which, following the occurrence of A, B, and C happen simultaneously, they will, in the absence of other cues, conclude that the structure is a common cause.

[1] In the causal Bayes net framework a causal model is defined in terms of a causal graph that represents the structural relations between variables in the model, and a set of parameters that determine the probability of each variable conditional on its direct causes (Pearl, 2000).

Temporal Information Versus Statistical Information

Experimental findings suggest that both statistical information and temporal information can in principle be used to make causal structure judgments. However, these are very different types of information, and moreover exploiting them is likely to differentially load on information processing resources. Bayesian approaches to causal learning have typically focused on how statistical information can be used to discriminate between causal structures (though see Rottman et al., 2014). However, Bayesian models of causal learning are normative rather than psychological accounts, and, as we have already pointed out, in practice the use of probabilistic information may be computationally very challenging and likely to place considerable demands on short-term memory (STM) and attention (Fernbach & Sloman, 2009; Meder, Gerstenberg, Hagmayer, & Waldmann, 2010; Rottman & Keil, 2012). As Fernbach and Sloman point out, there are 25 possible acyclic structures for even a simple three-variable system; it is psychologically implausible that participants in a causal structure learning task hold in mind all of these structures and make complex calculations about conditional probabilities in order to discriminate between them.

Fernbach and Sloman argue that participants may instead use what they term a structurally local computation strategy: participants may focus on confirming or disconfirming individual causal links in a structure, which they then combine together to infer a structure (for a related suggestion, see Waldmann, Cheng, Hagmayer, & Blaisdell, 2008). As they characterize it, this sort of local computation does not require that participants bear in mind multiple hypotheses and calculate conditional probability information extracted from aggregating information over a variety of observations. However, if in assessing evidence for individual causal links, participants relied only on processing some sort of statistical information (e.g., whether A covaries with B and whether B covaries with C), they would still need to piece together information that they had acquired about the different links from a number of different observations. That is, even if participants are using statistical information only to make local computations, it may still be necessary to combine information gleaned from different observations, making demands on processing resources such as working memory. In Fernbach and Sloman's (2009) terms, if participants relied on statistical information to make *structurally* local computations, such computations may still not be *temporally* local in that they would require putting together information aggregated across time.

By contrast, temporal cues to the whole causal structure, rather than just separate individual links within the structure, may be available in just a single observation. In the absence of any other cues to causation, participants are likely to draw a conclusion about causal structure based on the single observation that A, B, and C happen sequentially, or that, following A, events B and C happen simultaneously (Burns & McCormack, 2009). Here, we will refer to this type of temporal information, which may be available when a causal system involving three or more nodes operates, as *temporal pattern* information. Participants seem to use "rules of thumb" or heuristics when they observe such temporal patterns, inferring a causal chain in the former instance and a common cause in the latter case. This sort of heuristic, like other types of simple heuristics (Gigerenzer & Goldstein, 1996), places minimal demands on processing resources as participants do not have to track statistical information across observations or even combine information obtained from different observations.

In causal structure learning, the temporal pattern of events could be described, using the terminology introduced by White (2014), as a "singular" cue to causal structure: one that can be detected in single instances as opposed to one that necessarily requires multiple observations. White argues that a key set of singular cues (including temporal ones) plays a fundamental role in shaping hypotheses about causation, and that, at best, statistical cues such as covariation are only used subsequently in hypothesis-testing. Moreover, he argues that hypotheses set up as a result of singular cues may lead to statistical information being ignored or misinterpreted. Two studies relevant to this issue are those of Lagnado and Sloman (2004, 2006), which examined whether participants were likely to use simple temporal heuristics in their causal structure judgments, and the impact that this had on their use of statistical information. Lagnado and Sloman (2006) were interested in whether participants would set aside temporal heuristics when it is appropriate to do so and focus instead on statistical information. Use of temporal order cues is inappropriate if causal order does not in fact match temporal order; nevertheless Lagnado and Sloman (2006) demonstrated that even when participants were aware that these two types of order might not match, they relied heavily on temporal order in making causal structure judgments. Lagnado and Sloman's (2006) findings suggest that, consistent with White's (2014) claims, temporal cues might be more influential than statistical information, and lead to the latter information being underweighted or misused.

Note, though, that based on what has been discussed so far, we can distinguish between two possible reasons why participants might preferentially rely on the temporal pattern of events in a causal structure learning task. First, it may be that this type of singular cue is, as White argues, one of a set of fundamental cues that underpins causal judgments and will for this reason be accorded more weight than statistical information. Alternatively, it may be that using statistical information places considerably more demands on information processing resources than using temporal pattern information (as emphasized by Fernbach & Sloman, 2009). White's argument about the inherent importance of singular cues is based on his particular view of the origins of causal cognition, which he believes stems from our actions on the world. Temporal singular cues themselves are assumed to be grounded in the fact that our actions have certain temporal relationships to their outcomes (see White, 2014, for more detail). White suggests that even if people acquire more sophisticated abilities to process statistical information, they retain a "pervasive bias" (p. 68) to prefer singular cues. We return to a more in-depth discussion of what might underpin a bias toward temporal cues in the General Discussion, but regardless of its basis, an *inherent bias account* suggests a simple hypothesis: that participants will generally prefer temporal pattern cues over statistical information and will tend to base their causal structure judgments on them.

We can distinguish between the *inherent bias account* and the suggestion that temporal pattern information might be simply easier to use due to the information processing demands of using statistical information. We will refer to this second view as an *information processing account* of the use of temporal over statistical information. If it is correct, Lagnado and Sloman's results might be explained not in terms of a bias for temporal information but in terms of the complex nature of the task. The learning task that participants faced in Lagnado and Sloman's (2006) study was challenging, with participants having to bear in mind the instruction provided at the start of the experiment that temporal information was not reliable. Moreover, not only were the causal systems probabilistic (i.e., effects did not reliably follow their causes), but the range of possible hypotheses about causal structure was not constrained in the task. Participants did not select between a restricted set of possible structures but decided on whether there was a causal link (and in which direction) between each node in the structure; in Experiment 1 and some trials in Experiment 2 this was a four-node structure meaning that there were very many possible structures. This may have made it more likely that participants defaulted to relying on temporal heuristics.

What these considerations suggest is that it is important to examine whether participants default to using temporal heuristics even under circumstances in which the information processing demands of using statistical information are considerably reduced. If they did, then it would suggest that perhaps there is an inherent bias toward using such temporal information; we would then need to consider the basis of such a bias. In our experiments in the current article, similar to Lagnado and Sloman (2004, 2006), we asked participants to make causal structure judgments. However, we simplified the learning task considerably by narrowing the range of hypotheses to just three: in all the experiments participants had to decide whether a three-variable system was either a common cause one, B←A→C, an A→B→C causal chain, or an A→C→B causal chain. We then systematically examined participants' use of temporal pattern and/or statistical information across three experiments; these experiments differed primarily in the means by which participants were provided with statistical information. We were interested in whether participants would be more likely to use simple temporal heuristics that could be based on a single observation of the system than any type of statistical information. We examined whether this was the case even under circumstances in which we made it relatively easy for participants to make use statistical information by making structurally local computations (i.e., pairwise computations). In Experiments 2 and 3, we optimized participants' opportunities to confirm or disconfirm individual causal links between nodes in the system on the basis of dependency information by allowing them to separately observe the effects of intervening on each node in the system on each of the other nodes in a fully deterministic system.

One important way in which these experiments differed from those previously reported was that we also included groups of children as well as adults. Setting aside the intrinsic interest in mapping the developmental profile of causal structure learning, from the perspective of comparing the use of statistical information versus temporal pattern information a key advantage of including groups of children is that it provides a straightforward way of varying the information processing resources of participants. If use of temporal heuristics makes minimal information processing demands compared with using statistical information (even information used in local computations about pairwise causal links), we would expect young children to rely predominantly on temporal information. Thus, there should be developmental changes in the extent to which participants are able to exploit statistical information in comparison with temporal pattern information. However, if there is an inherent bias toward the use of temporal information, then on White's (2014) account we might also expect children to be particularly reliant on such information. White argues that such a bias stems from the developmental origins of causal inference, and use of statistical information is something that is mastered only later in development. Thus, both an inherent bias account and an information-processing account of a preference for temporal information predict that children will predominantly rely on temporal pattern rather than statistical information in causal structure learning. Of course, the inherent bias account and the information processing account need not be seen as being in opposition to each other: It could be the case that both accounts are correct and may together explain preference for temporal information. However, the inherent bias account makes an additional prediction over and above the information processing account, which is that it predicts that participants with ample information processing resources (i.e., adults) will tend to rely on temporal information even under circumstances in which the information processing demands of using statistical information are minimized.

Children's Causal Structure Learning

We have suggested that even using statistical information in a structurally local way to learn pairwise causal links might make some demands on information processing resources, because it requires some sort of integration across observations. Children might be expected to have difficulties with such integration, due to their limited processing resources. There is a well-established tradition of cognitive developmental research that emphasizes the impact that limited resources have on children's ability to integrate information in their learning (e.g., Halford, 1984, 1993; Halford & Andrews, 2006). Recent developmental studies have not focused on the impact that these limited resources may have on children's causal learning; instead such studies have emphasized the similarity of children's learning to that of an idealized Bayesian learner (for review, see Gopnik, 2012; Gopnik et al., 2004). However, some studies have suggested that there are qualitative changes in young children's causal learning, and that these are in part attributable to changes in processing resources (McCormack, Butterfill, Hoerl, & Burns, 2009; McCormack, Simms, McGourty, & Beckers, 2013; Simms, McCormack, & Beckers, 2012), with McCormack, Simms, McGourty, and Beckers' (2013) study demonstrating that the developmental emergence of cue competition effects in a causal learning task is linked to children's working memory capacity. However, these studies specifically focused on the cue competition effect of blocking, and did not examine causal structure learning.

In fact, we know relatively little about children's ability to use temporal and/or statistical information to discriminate between, for example, common cause versus causal chain structures. We know that children aged from at least 6 years can use simple temporal heuristics of the sort described above to make causal structure inferences (Burns & McCormack, 2009; Frosch, et al., 2012). However, we do not know whether young children can use any type of statistical information to make such discriminations, let alone the status of such information relative to temporal heuristics. In one relevant study by Schulz, Gopnik, and Glymour (2007), preschoolers were shown pairs of gears, B and C, housed in a box that was operated with a switch, and were shown whether or not the operation of one gear depended on that of the other gear by means of the experimenter making interventions on each gear in turn (removing it from the box). For example, in one trial type children saw the experimenter demonstrate that when gear B was removed gear C did not operate, but when gear C was removed, gear B did still operate. In another type of trial, children were shown that B operated even without C and that C operated even without B. Schulz et al. (2007) asked children to select between anthropomorphized pictures that depicted the relationships between the pairs of gears. For the first trial type described, selecting a picture in which cartoon hands emerging from B pushed C, but not vice versa, was the correct answer, whereas for the second type the correct answer was to select a picture in which neither gear pushed each other. Four- to 5-year-olds appropriately selected pictures following the demonstrations in which the experimenter made interventions on each of the two gears in turn.

The authors argued that their findings provided evidence that children of this age can discriminate between causal chain and common cause structures on the basis of interventions that they observe. Although this study does indeed demonstrate that preschoolers can rapidly learn whether the operation of pairs of components is interdependent (see also Rottman et al., 2014), it is much less clear whether it shows that children of this age can readily use information from interventions to discriminate between three-component causal structures, such as those depicted in Figure 1. Although there was a third component in the system (an on-off switch), children were not required to represent how its relationship to the gears differed between the different trials, and the switch itself was not depicted in the diagrams used to give their responses. Rather, children's responses could simply have reflected their knowledge of the relationships between B and C. That is, they need not have possessed representations that reflected the sort of causal structures depicted in Figure 1; to put the point in terms of the relationships depicted in this figure, the judgments required only knowing whether there was an arrow between B/C and in which direction.

We have argued above that even if participants are able to learn individual pairwise causal links between nodes in a causal system (i.e., use a structurally local approach), integrating this information to construct a causal structure may make further demands on information processing resources. This consideration is particularly relevant in the context of developmental studies, given the well-documented constraints in young children's processing resources. Thus, although Schulz et al.'s (2007) study may have demonstrated that young children can learn a pairwise dependency, this demonstration does not allow us to conclude that they can use statistical information (such as that provided by interventions) to learn the structure of a three-variable system.

Sobel and Sommerville (2009) conducted a study that followed up that of Schulz et al. (2007). In their study, children saw a set of three colored lights, A, B, and C and were told that some lights cause other lights to work. They were initially shown that when A was switched on, the other two lights also illuminated. They were then shown an intervention on light B (it was covered with a cup), which they were told prevented it making any other lights work. Following this intervention, in one trial type (common cause), C still operated when A was switched on, whereas in the other trial type (ABC causal chain), C did not operate when A was switched on. After observing each trial type, 4-year-olds had to choose between one of two sets of pictures: One set comprised of a picture of A causing B and a picture of A causing C; the other set comprised of a picture of A causing B and a picture of B causing C. Thus, although the responses depicted all three variables, rather than two as in Schulz et al.'s (2007) study, arguably children still had to consider only the pairwise relationships between the variables (specifically, whether there was a causal link between B and C), rather than the overall causal structure. A similar issue arises with regard to interpreting the findings of a further study by Sobel and Sommerville (2010) in which children were only questioned about each pairwise link separately.

Other aspects of Sobel and Sommerville's (2009) study also raise doubt over whether children in their study properly discriminated between causal chain and common cause structure. Children initially saw that switching on A made B and C also turn on. Even if they ignored the effects of the disabling intervention on B, they would be expected to choose the set of pictures in which A makes B go and A makes C go, meaning that responses in the common cause trial do not inform us whether or not children were using the information from the intervention on B. Of more interest is whether children were successful on the causal chain trial. In that trial, 14 of their sample of 20 children chose the set of pictures in which A made B go and B made C go, which just fails to reach conventional levels of significance (given 50% chance of choosing the correct set). We note that participants could have chosen the correct picture in the causal chain trial not because they understood that the structure was a causal chain but because they had just seen a demonstration in which on pressing A, C did not light up. This could have led them to conclude that A does not cause C, and hence they could have decided against the set of pictures in which A was depicted as causing C.

The more general point is that in deciding if children grasp that a structure is an ABC causal chain, it is not sufficient to demonstrate that they believe there is a causal link between A and B, and, separately, believe there is a causal link between B and C (to use an, albeit imperfect, analogy, understanding $X > Y$ and $Y > Z$ does not mean that one necessarily can make the transitive inference that $X > Z$). Understanding the structure of the system involves putting these two pieces of information together appropriately to reflect the fact that causality is propagated through the chain. In summary, looking across the small number of existing studies of children's causal structure learning, it is difficult to know whether young children are capable of learning causal structure on the basis of statistical information, particularly if we bear in mind the distinction between learning structurally local pairwise links and integrating such links to form a representation of causal structure.

Current Studies

The three experiments reported here examined children's and adults' causal structure learning in tasks in which they were explicitly required to differentiate between common cause and causal chain structures. Our key aims were to examine (a) how competent participants are at using statistical information in causal structure learning in comparison with temporal pattern information, and (b) how this might vary with age. According to both an inherent bias and an information processing view of temporal information, we might expect to see a preference for temporal pattern information over statistical information that is particularly marked in children. Compared with those used in previous studies of adult causal structure learning, the tasks in the current study were relatively simple and did not require participants to keep in mind a cover story about the variables in the system. Across three experiments, children and adults observed a three-component mechanical system and were required to decide which of three structures it matched, ensuring participants needed only to consider a narrow range of hypotheses. The current study differed from Schulz et al.'s (2007) and Sobel and Sommerville's (2009) studies with children in that we used a wider age range of children: They focused on 4- to 5-year-olds, whereas our children ranged in age from 5 to 8 years. Unlike in those studies, we asked participants to make a judgment about causal structure using the sort of diagrams typically used in adult studies, in which the relationships between all three variables are depicted in a single structure.

Our three experiments differed in terms of the way statistical information was provided: participants either observed a system operating probabilistically (Experiment 1) or saw the results of interventions being carried out on the nodes in a deterministic system (Experiments 2 and 3). Table 1 provides an overview of the experiments and their conditions; conditions will be explained below. In Experiment 1, in which the causal system operated probabilistically, participants observed the system operating without any interventions taking place. However, the probabilistic information was sufficient to discriminate between the causal structures. In one condition, though, temporal pattern and statistical information were placed in opposition to each other, and of interest was whether participants would use temporal or statistical cues. In Experiments 2 and 3, statistical information was provided by the experimenter making interventions on a deterministic system. We used two different types of intervention. In Experiment 2, participants saw "generative interventions:" interventions in which the experimenter made each component in the system operate in turn, and demonstrated the effect that this had on other components in the system. For example, participants saw that operating A made both B and C go, and also that operating B made C go but operating C did not make B go. Such a pattern would suggest that the system was an ABC causal chain.

In Experiment 3, participants saw what we will term "prevent-then-generate" interventions. For these interventions, the experimenter disabled a component in the system and then demonstrated what effect this had on the operation of other components in the system. For example, the experimenter might disable B and show that C still operates when A is made to go, and then disable C and show that B operates when A is made to go. This pattern of observations suggested a common cause structure. Compared with most previous studies with adults, what was distinctive about Experiments 2 and 3 was that the system was fully deterministic and that each node was systematically intervened on in turn. These conditions should have maximized the opportunity for participants to observe the dependencies between each pair of nodes in the system, facilitating the sort of structurally local computations described by Fernbach and Sloman (2009). We were interested in whether children and adults used this statistical information and how easy they found it to use compared with temporal pattern information.

Experiment 1: Time and Statistical Information

The first study pitted temporal pattern and statistical information against one another in causal structure learning. In the congruent condition, temporal and statistical information were consistent with one another, with both suggesting the same causal structure. In the incongruent condition the temporal and statistical information were inconsistent with one another and suggested different causal structures (e.g., the conditional dependencies and independencies suggested a common cause, but the temporal pattern was sequential, suggesting a causal chain structure). We also included a statistical information only condition in which all events occurred simultaneously, to examine if participants would use statistical information alone to discriminate between the structures. We note that in the literature there is some doubt over whether even adults can reliably extract causal structure just by observing the operation of a probabilistic system (Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Rottman & Keil, 2012; Steyvers et al., 2003). However, as we have pointed out, the participants' task was simplified considerably in this study by telling participants in advance that the system could have one of only three causal structures, and by making the system only quasi-probabilistic (events never occurred spontaneously without their putative causes) and pseudorandom (participants were provided with a relatively small predetermined set of observations that would be sufficient to distinguish between structures).

Method

Participants. Children from three different school classes and one adult group participated in this study: 92 5- to 6-year-

Table 1
Summary of the Types of Information Available in Each Experiment

Exp.	Condition	Type of statistical information	Temporal pattern of all three events
Exp. 1	Congruent	Observation of covariation in probabilistic system	Sequential for causal chain; synchronous for common cause
	Incongruent	Observation of covariation in probabilistic system	Sequential for common cause; synchronous for causal chain
	Statistical only	Observation of covariation in probabilistic system	No delays
Exp. 2	Intervention no delay	Each node is operated separately to demonstrate its effects on other nodes	No delays
	Time only	None	Sequential for causal chain; synchronous for common cause
	Intervention plus time	Each node is operated separately to demonstrate its effects on other nodes	Sequential for causal chain; synchronous for common cause
Exp. 3	N/A	Each node disabled separately and then each other node operated	N/A (all three events never observed together)

olds (M = 77 months, *range* = 70–83 months); 69 6- to 7-year-olds (M = 88 months, *range* = 83−94 months); 64 7- to 9-year-olds (M = 100 months, *range* = 95–114 months); and 74 adults. Participants completed one of three conditions: the congruent condition (30 5- to 6-year-olds; 17 6- to 7-year-olds; 22 7- to 9-year-olds; and 21 adults); the incongruent condition (31 5- to 6-year-olds; 17 6- to 7-year-olds; 21 7- to 9-year-olds; and 22 adults); or the statistical information only condition (31 5- to 6-year-olds; 35 6- to 7-year-olds; 21 7- to 9-year-olds; and 31 adults). There were 164 females in total. The children were tested individually at four schools. The adults were psychology students who either received course credit or were paid for their participation.

Apparatus. A purpose built wooden box measuring 41 cm (long) × 32 cm (wide) × 20 cm (high), with an on-off switch at the front, was used. There were different colored lids for the box, each of which had three objects inserted on its surface that rotated on the horizontal plane when the experimenter activated a remote button (see Figure 2). Each lid had three predetermined locations for these objects that formed an equilateral triangle of sides 24 cm. Unbeknownst to participants, the operation of the box was controlled by a laptop computer hidden inside it. In the congruent condition, the components operated on three different temporal schedules according to the causal structure they were presenting. When a box displayed the causal structure shown in Figure 2a, the A component rotated for 1 s followed by a 0.5 s pause followed by the simultaneous rotation of components B and C which lasted for 1 s (synchronous schedule). When a box displayed the causal structure in Figure 2b, the A component rotated for one second followed by a 0.5-s pause followed by the 1-s rotation of component B, which was also followed by a 0.5-s pause and the 1-s rotation of component C (sequential schedule). The operation of the causal structure displayed in Figure 2c was similar to that of 2b with the difference that component C rotated prior to component B. The colors and shapes of the components were varied across participants and causal structures. Photographs of the box depicting the three possible causal structures were used at test (as in Figure 2), overlaid with pictures of hands to indicate causal links (following Schulz, Gopnik, & Glymour, 2007).

Design. We employed a between-participants design, with participants assigned to one of the three conditions (data from the statistical information only condition were collected at a later time point than the other two conditions but from a very similar population). Each participant was shown either two congruent trials where timing and covariation information were consistent with the same causal structure (one common cause and one causal chain—either ABC or ACB—each consisting of 18 demonstrations; Table 2), two incongruent trials, or two statistical information only trials in which all events occurred simultaneously. The order in which participants were shown each demonstration was randomized for each participant, and the order in which participants received each trial type was counterbalanced. In the incongruent condition, when the statistical information suggested a causal chain, the temporal schedule was a synchronous one, and when the statistical information indicated a common cause, the temporal schedule was a sequential one. The statistical information only condition contained no temporal information, in the sense that all components rotated simultaneously.

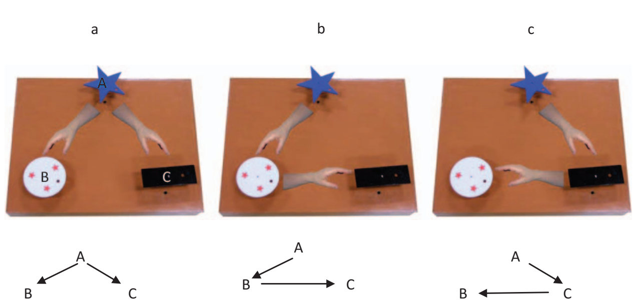

Figure 2. The three different causal structures displayed by the boxes: (a) Common cause (A causes both B and C); (b) Causal chain 1 (A causes B and B causes C); (c) Causal chain 2 (A causes C and C causes B). See the online article for the color version of this figure.

Materials. Three causal models were used to generate the patterns of statistical data for each condition:

Model 1: Common cause (B←A→C)

Model 2: ABC Chain (A→B→C)

Model 3: ACB Chain (A→C→B)

Each model implies a different set of conditional independence relations and thus is in a separate Markov-equivalence class. For Model 1, B is independent of C, conditional on A; for Model 2, A is independent of C, conditional on B; for Model 3, A is independent of B, conditional on C. This means that each model is in principle distinguishable on the basis of observational data alone. All models have the same parameters for the strengths of individual links ($s = 0.66$) and there are no spontaneous "uncaused" effects ($e = 0$).

Each model was used to generate a representative pattern of data to be presented to participants. The event frequencies for each model's dataset are shown in Table 2; each dataset consists of 18 trials. Each dataset is fully representative of the underlying generating structure. Thus, in the dataset for Model 1, P(B moves |A moves) = 0.66, P(C moves |A moves) = 0.66; for Model 2, P(B moves |A moves) = 0.66, P(C moves |B moves) = 0.66; for Model 3, P(C moves |A moves) = 0.66, P(B moves |C moves) = 0.66. Note that the models were semideterministic insofar as there were no spontaneous effects. Therefore, for Model 2, P(C moves| B does not move) = 0, and for Model 3, P(B moves | C does not move) = 0. This means that Model 2 is refuted by a pattern in which A and C move but B does not move, and Model 3 is refuted by a pattern in which A and B move but C does not. There is no corresponding refuting pattern for Model 1.

How discriminable are the three different models based on the datasets presented to participants? This question in part depends on the learning procedures used. As noted in the introduction, there are two main computational approaches to structure learning: Bayesian and constraint-based. We will focus on a Bayesian approach, because it is more readily applied to the current learning context. Although the three structures are in principle discriminable via constraint-based methods, because each has a distinct set of conditional independencies, these methods typically require large datasets to test for conditional independence (Spirtes, Glymour & Scheines, 1993). Moreover, constraint-based methods can struggle with semideterministic systems, as used in the current experiments (Cooper, 1999).

Table 2
Event Pattern Frequencies and Likelihoods for the Three Possible Causal Structures in Experiment 1

	Causal structure					
	Common cause		ABC chain		ACB chain	
	MODEL 1		MODEL 2		MODEL 3	
Event pattern (components moving)	f	l	f	l	f	l
A, B, and C	8	0.44	8	0.44	8	0.44
A only	2	0.11	6	0.33	6	0.33
A and B only	4	0.22	4	0.22	0	0
A and C only	4	0.22	0	0	4	0.22

Note. f = event pattern frequency; l = likelihood of event pattern given model (f/18).

Bayesian Analysis of Data Sets

A standard Bayesian approach assigns prior probabilities to all candidate models, and then updates these given the data using Bayes' rule to yield posterior probabilities for each model:

$$P(model_i \,|\, data) = \frac{P(data \,|\, model_i) \cdot P(model_i)}{\sum_i P(data \,|\, model_i) \cdot P(model_i)} \quad (1)$$

Here i = 1, 2, or 3, for the three different causal models. Each model has its corresponding dataset (data1, data2, data3). To compute posteriors for each model we require likelihoods for each specific pattern of data given each model. The objective likelihoods for each data pattern given each model are shown in Table 2. For example, P(A,B,C| Model 1) = 0.44; P(A, not-B, not-C| Model 1) = 0.11; and so forth.

For our first analysis we assume an ideal Bayesian learner that knows the correct model parameters, that is, that each link has strength s = 0.66 and there are no spontaneous effects, e = 0. We also assume a uniform prior probability distribution over the three models, so each model has a prior of one third. Bayesian computations are carried out for each dataset separately (data1, data2, data3). The results of this analysis are presented in the first column in Table 3. They clearly show that the models are discriminable by an ideal Bayesian learner; the posteriors for the correct models are one or close to one in all cases. However, unlike the Bayesian model, the participants in this experiment were not given the parameters of the models, and would have to make assumptions or learn them during the task. One formal approach is to use maximum likelihood estimation, and indeed the datasets would support accurate estimates for the strengths of the links if participants infer the correct model. But the simultaneous learning of structure and parameters is a more complicated task (Griffiths & Tenenbaum, 2009; Heckerman, 1998). Rather than explore more complex Bayesian models here, we will show that for a range of plausible parameters that participants might assume, the three models are still strongly discriminable.

Two key model parameters are varied: the strength of the links and the level of spontaneous "uncaused" effects.

Strength of links. We included a range of parameter values for the link strengths, with the constraint that within-model links have the same strength. As noted above, participants are exposed to a representative dataset and thus their strength estimates should not be too far from the correct value. We used three link strength parameters: s = 0.5, 0.66, and 0.8.

Spontaneous effects. Given the instructions and the mechanical set-up it seems reasonable for participants to assume that spontaneous effects are very unlikely. We included a range of parameter values for spontaneous effects (e = 0, e = 0.15 or 0.30) corresponding to several assumptions that participants might make, with the constraint that spontaneous effect rates were the same for within-model effects.

Table 3 shows the posteriors for each model for combinations of different link strengths and spontaneous effect rates. For all combinations the correct model is clearly discriminable. More generally, additional tests confirm that the correct model is the most probable even with extremely high levels of spontaneous effects (e.g., 0.8 or 0.9). The results also show that the distinguishability of the common cause model decreases as the level of spontaneous effects increases, but the chain models are much less sensitive. Overall, these analyses suggest that the correct models are dis-

Table 3
Discriminability of Structures Given the Statistical Data in Experiment 1

	Link strength								
	0.67			0.50			0.80		
Posterior of model	e = 0	e = .15	e = .30	e = 0	e = .15	e = .30	e = 0	e = .15	e = .30
P(model 1 | data 1)	1	.967	.706	1	.964	.734	1	.949	.575
P(model 2 | data 1)	0	.017	.147	0	.018	.133	0	.025	.213
P(model 3 | data 1)	0	.017	.147	0	.018	.133	0	.025	.213
P(model 1 | data 2)	.001	.002	.002	.015	.015	.015	0	.000	.000
P(model 2 | data 2)	.999	.997	.975	.985	.980	.943	1	.999	.985
P(model 3 | data 2)	0	.002	.023	0	.005	.043	0	.001	.015
P(model 1 | data 3)	.001	.002	.002	.015	.015	.015	0	.000	.000
P(model 2 | data 3)	0	.002	.023	0	.005	.043	0	.001	.015
P(model 3 | data 3)	.999	.997	.975	.985	.980	.943	1	.999	.985

Note. Posterior probabilities for each structure (model 1 = common cause; model 2 = ABC chain; model 3 = ACB chain) given each dataset presented to participants (data1, data2, data3). Bayesian updating was used to infer the posteriors using a uniform prior and varying model parameters for strength of link (e = .67, .50 or .80) and probability of spontaneous effects (e = 0, .15 or .3).

criminable given the datasets given to participants, even without temporal order information.

Procedure. The children were tested individually and were first introduced to the box and asked to name the color of each of the components. They were then shown three pictures, such as the ones in Figure 2, which were described to them as showing how the box might work. The pictures with the superimposed hands were explained to them (e.g., for Figure 2a the experimenter said "In this picture the blue one makes both the black one and the white one go, and the hands show that."). Children were then asked comprehension questions that required identifying each of the three pictures that had just been shown to them, for example, "Can you show me the picture where the blue one makes both the white one and the black one go?" When children made errors on the comprehension questions the pictures were described to them again and they were asked the comprehension questions again.

On completion of the comprehension questions the children's attention was drawn to the on and off switch at the front of the box and they were asked whether the box was switched on or off (it was always off). They were then told "In a moment I am going to switch the box on and I want you to watch carefully what happens. Remember, you've got to figure out which of these pictures shows how the box goes. The [B] and [C] ones move but they don't always work. So let's switch the box on now." The children then observed the appropriate 18 demonstrations of the box operating (see Table 2). After observing the 18 demonstrations, in all conditions they were asked to identify which of the three pictures "shows how the box really works." The children then completed a maze task for a few minutes which served as a filler task before moving on to the second trial. The lid of the box was replaced with a different colored lid and new components were introduced. Children were told that the new box may work the same or may work differently to the one they had already seen.

The procedure for the adults was very similar with the following exceptions. Some of the adults were tested in pairs rather than individually, in which case responses were given in writing rather than verbally, and the short filler task for the adults consisted of an unrelated verbal reasoning task. Adults were not asked the initial comprehension questions.

Results

Thirty-four of the children (23 of the 5- to 6-year-olds: five from the congruent condition, seven from the incongruent condition, and 11 from the statistical information only condition; six of the 6- to 7-year-olds: two from the congruent condition, and four from the statistical information only condition and five of the 7- to 8-year-olds: one from the congruent condition, one from the incongruent condition, and three from the statistical information only condition) did not pass the comprehension trials first time. We made the conservative decision to exclude these children from the analyses because of the possibility that children who needed multiple comprehension trials may have had difficulty understanding the nature of the task. However, in this and in subsequent experiments, the same qualitative pattern of findings is obtained if all children are included.

As can be seen in Figure 3, accuracy in the congruent condition was high across all of the age groups, with all groups choosing the answer consistent with probabilistic information significantly

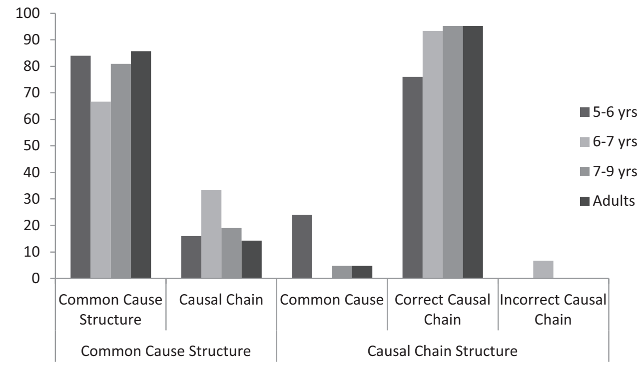

Figure 3. The percentage of causal model choices for each structure in the congruent condition in Experiment 1 as a function of age group. Statistical information and temporal cues were congruent in this condition.

more often than chance for all causal structures (binomial test $p <$ .05, assuming participants would be correct one third of the time by guessing). Figure 4 shows participants' responses in the incongruent condition for each of the two causal structures; for the common cause structure, the temporal cues indicated a causal chain, whereas for the causal chain structure, the temporal cues indicated a common cause structure. It can be seen from the figure that relatively few participants chose the causal structures that were indicated by the statistical information. Participants in all groups chose the structure consistent with the temporal schedule significantly more often than chance for both causal structures (binomial test, $p < .05$), with even adults only infrequently choosing the structure consistent with the statistical information. In the statistical information only condition (see Figure 5), although performance improved with age, all groups, with the exception of the youngest group, chose the causal structure consistent with statistical information more often than chance for the common cause structure (binomial test, $p < .05$). However, for the causal chain structures, only the adult group chose the correct structure significantly more often than chance (binomial test, $p < .05$, assuming participants would choose the correct structure one third of the time by guessing). This was due to children's tendency to choose the common cause structure in these trials as well as in the trials for which this was the correct answer (57.3% of children's choices were common cause in causal chain trials). It can be seen from Figure 5 that even adults, who were above chance on this trial, were as likely to incorrectly give common cause judgments for the causal chain as to correctly give causal chain judgments.

Discussion

Our results paint a clear picture of the importance of temporal pattern information in causal structure learning. In the incongruent condition, temporal rather than statistical information guided causal structure identification. This was true even for adults, although adults were above chance in their use of statistical information in the statistical information only condition. These results confirm Lagnado and Sloman's (2004, 2006) suggestion that temporal cues will be weighted more heavily than statistical cues, but extend them by indicating that this is the case even if the task is simplified considerably. Although some of Lagnado and Sloman's

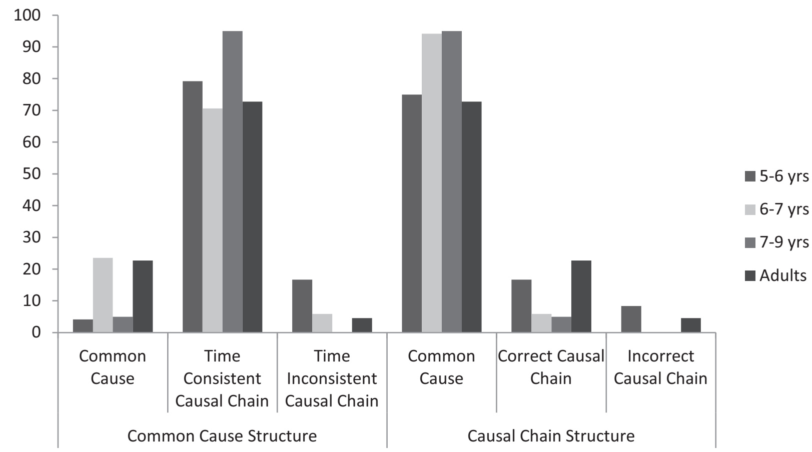

Figure 4. The percentage of causal model choices for each structure in the incongruent condition in Experiment 1 as a function of age. When statistical information indicated a common cause structure, the temporal cues suggested the causal chain labeled as time consistent; when statistical information indicated a causal chain structure, the temporal cues suggested a common cause structure.

experiments involved making similar judgments about three-variable causal systems, their design was much more complex, involving situations that had to be contrived so that the temporal order in which participants found out about events did not necessarily match the order in which the events actually occurred. Even adults may have struggled with this decoupling of experienced temporal order from actual temporal order. By contrast, in the current paradigm, temporal information was simply manipulated within a physical system in such a way that different causal structures remained plausible without a complex cover-story.

It is important to point out that although we describe the temporal pattern information as being incongruent with probabilistic information in the incongruent condition, nevertheless it was still physically possible for the causal structure consistent with probabilistic information to obtain. For example, if participants see a series of events ABC with a sequential schedule, it is perfectly possible that the true causal structure is one in which A is a common cause of B and C, but that the delay between A and C's occurrence differs from the delay between A and B's occurrence.

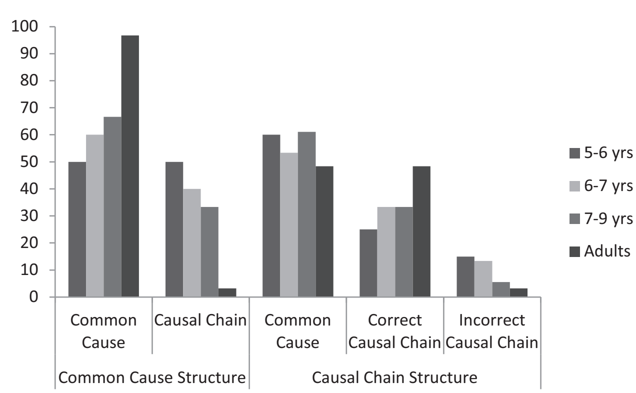

Figure 5. The percentage of causal model choices for each structure in the statistical information only condition in Experiment 1. In this condition all components moved simultaneously.

Although this state of affairs was physically possible, our results indicate that participants nevertheless overwhelmingly chose to use the structure suggested by the temporal pattern of events rather than the statistical information. Thus, participants judged that they were viewing a causal chain ABC when the pattern was sequential, even though they saw C occurring in the absence of B a number of times.

Adults were able to discriminate between causal structures on the basis of statistical information alone, but even this group only used statistical information around half of the time when making judgments about the causal chains in the statistical information only condition. Although we have termed this condition statistical information only, the events that occurred did have a temporal pattern, insofar as they all happened simultaneously. When all events occurred simultaneously, even the oldest children preferred to choose a common cause structure regardless of the pattern of covariation. This finding confirms the status of temporal pattern information relative to statistical information, because it again suggests that if events happen simultaneously, they are likely to be viewed as the result of a common cause rather than forming a causal chain. It may be that when faced with a mechanical system such as the one used in our study, both children and adults take temporal cues to be more revealing of underlying mechanisms; we return to this issue in the General Discussion.

Experiment 2: Time and Interventions

In Experiment 1, causal relationships between events were probabilistic to allow participants to observe the system operating and then use statistical information to discriminate between causal structures. We found that children were unable to use this information. However, most previous studies of children's causal learning have used fully deterministic causal relationships. Indeed, there is evidence to suggest that children find it difficult to deal with probabilistic causal relationships and are likely to work with a default assumption that systems are deterministic (Schulz & Sommerville, 2006). Moreover, even adults were far from perfect at using probabilistic information in Experiment 1, consistent with previous research that has suggested that simply observing the operation of a probabilistic system may not be sufficient for good causal structure learning (Fernbach & Sloman, 2009; Steyvers et al., 2003). Indeed, it could be argued that Experiment 1 did not provide a completely fair test of the relative weighting accorded to temporal versus statistical information precisely because recruiting statistical cues in this experiment required extracting information provided by the stochastic nature of the system. By contrast, the temporal pattern cues were deterministic in the sense that they were consistent and completely uniform across trials.

For these reasons, in our second experiment, we used systems that were deterministic and instead demonstrated the effects of interventions on components in the system in order to provide participants with information about the dependencies and conditional independencies between events. The interventions used in this experiment were what we have termed generative: the experimenter moved a particular component in the system and demonstrated the effect that this intervention had on the operation of the other components. The experimenter repeated this for each node in the system. These circumstances should help participants to use dependency information to establish whether or not there is a causal link between each pair of nodes; thus, causal structure learning should be possible even if participants only use statistical information to make local computations.

In this experiment, we again varied whether this statistical information was accompanied by temporal information, by using three different conditions that differed in terms of whether intervention and/or temporal cues were provided. In the intervention no delay condition, participants saw generative interventions (i.e., interventions in which the experimenter selectively made components operate) on the same three-component system used in Experiment 1. In this condition, the experimenter made each of the components operate in turn and allowed participants to observe the effect that this had on the other components in the system. No temporal order cues were provided, in that all events happened simultaneously (moving A also made B and C move at the same time). In this sense, the methodology was similar to that of Schulz et al. (2007) and Sobel and Sommerville (2009), who also used simultaneous events. The generative interventions themselves provided sufficient information to discriminate between causal chain and common cause structures (e.g., moving B will make C go if the structure is an ABC causal chain but not if it is a common cause structure; similarly, moving B will not make C go if A is a common cause of B and C).

The second condition was a time only condition, in which participants were provided with temporal pattern information but not intervention information: When the experimenter switched on the box the components operated without interventions, but the temporal schedule of the operation of the components was either a synchronous one (A, followed by a delay, then B and C simultaneously) or a sequential one (A, followed by a delay, then B, then a further delay, then C for an ABC causal chain). The findings from the congruent condition in Experiment 1 indicate that participants will assume that the former system is a common cause, and the latter a causal chain. The third condition was an intervention plus time condition. In this condition, participants observed the same interventions as in the intervention no delay condition, but they were also provided with consistent temporal information. So, for example, if the structure was an ABC causal chain, when A was intervened on by the experimenter, B moved after a short delay, and then C moved after a further delay (a sequential temporal schedule).

Different predictions can be made depending on whether it is assumed participants can use statistical information from interventions and/or temporal pattern information in their causal structure learning. If participants can learn from observing interventions because this provides conditional dependency information, they should be able to discriminate between causal structures even in the intervention no delay condition. If, however, participants have difficulty using statistical information (as the findings of Experiment 1 suggest), then they may be successful only in the other two conditions in which temporal cues are provided. A third possibility is that participants may use the statistical information provided from observing interventions additively along with temporal information, in which case the best level of performance should be in the intervention plus time condition.

Participants

One-hundred and 40 children (68 girls) and 82 adults took part in the study. Children were from two different school years: 6- to 7-year-olds (M = 80 months; *range* = 74–86 months) and 7- to 8-year-olds (M = 92 months; *range* = 86–98 months). Participants were assigned to one of three conditions (intervention no delay, N = 21 6- to 7-year-olds, 25 7- to 8-year-olds, and 26 adults; intervention plus time, N = 24 6- to 7-year-olds, 25 7- to 8-year-olds, and 27 adults; and time only, N = 20 6- to 7-year-olds, 25 7- to 8-year-olds, and 29 adults). Children were recruited and tested individually in their schools. Adults were recruited through a university participant database. They completed the task in a laboratory along with some other unrelated cognitive tasks and received a small payment for their participation.

Materials

The materials were identical to those used in Experiment 1.

Design

We employed a between-participants design, with participants assigned to one of the three conditions in which their task was to discriminate between a common cause structure, an ABC causal chain, and an ACB causal chain. Although in Experiment 1, participants saw only one type of causal chain (either ABC or ACB), we included both of these chain types for all participants in Experiment 2. Thus, participants received three trials in total. Each participant was shown either three time only trials where they observed the operation of the box without the experimenter's interventions (the common cause and the two causal chains, each consisting of three demonstrations), three intervention no delay trials in which the experimenter intervened on each component twice to demonstrate how it affected the other components but no temporal information was provided (in the sense that components rotated simultaneously), or three intervention plus time trials in which the experimenter intervened as in the intervention only condition but the interventions were accompanied by the same temporal delays that were observed in the time only trials. The order in which participants were shown the interventions was randomized for each participant, and the order in which participants received each trial type was counterbalanced.

Procedure

Children were introduced to the box and answered comprehension questions as in Experiment 1. On completion of the comprehension questions the children's attention was drawn to the on and off switch at the front of the box and they were asked whether the box was switched on or off (it was always off). In the intervention plus time and the intervention no delay condition, children were then told "In a moment I am going to switch the box on and I want you to watch carefully what happens. For each shape I am going to show you what happens to the other shapes when I spin it. Remember, you've got to figure out which of these pictures shows how the box goes. So let's switch the box on now." The children then observed the appropriate demonstrations of the box operating, with the experimenter intervening on each component separately by rotating it. These interventions were carried out twice for each component. Thus, for the common cause structure, children saw that operating A made both B and C go, but that operating B or C had no effect on the other components. For both the ABC and ACB causal chain, children saw that operating A made both B and C go; for the ABC chain they also saw that operating B made C go, but operating C did not make B go, whereas the reverse was true for the ACB causal chain.

In the time only condition, children observed three demonstrations of the box components operating once the box had been switched on (as in Experiment 1), and the experimenter did not intervene to selectively operate any of the box components. After observing the demonstrations they were asked to identify which of the three pictures shown in Figure 2 "shows how the box really works." The children then completed an unrelated filler task (completing a maze) for a few minutes before moving on to the second trial. The lid of the box was replaced with a different colored lid and new components were introduced. Children were told that the new box may work the same or may work differently to the one they had already seen. They then completed a second unrelated filler task before being introduced to a third box.

The procedure for the adults was very similar with the following exceptions. Some of the adults were tested in pairs rather than individually, in which case responses were given in writing rather than verbally. Adults were not asked the initial comprehension questions and did not complete filler tasks.

Results

Seventeen of the children did not pass the comprehension trials first time (12 of the 6- to 7-year-olds: four from the intervention no delay condition, five from the intervention plus time condition, and three from the time only condition; and five of the 7- to 8-year-olds: one from the intervention no delay condition, two from the intervention plus time condition, and two from the time only condition); as in Experiment 1 we conservatively excluded them from the analyses although the qualitative pattern of the findings is unchanged if all children are included. The percentage of participants in each group who chose each response for the common cause and the causal chain structures is shown in Figure 6; Figure 6a shows performance for the intervention only condition, Figure 6b for the intervention plus time condition, and Figure 6c shows performance for the time only condition. It can be seen from the figures that performance was much better for the causal chains in the conditions that included a time cue than in the intervention no delay condition. Notably, children were likely to erroneously give a common cause response for the causal chain structures in the intervention only condition, and even adults were less likely to give a correct response for causal chains in this condition than in the other two conditions.

An initial ANOVA on the number of correct causal structure choices (scores ranging from 0 to 3, and in the time only condition choosing the common cause structure was categorized for the purposes of these analyses as correct for the synchronous schedule and choosing the appropriate causal chain as correct for the sequential schedules) with between-subjects factors of age group and condition found a main effect of age, $F(2, 196) = 11.50$, $p < .001$, $\eta_p^2 = .11$, and a main effect of condition, $F(2, 196) = 62.32$, $p < .001$, $\eta_p^2 = .39$. The interaction between age and condition was not significant, $F(4, 196) = 1.68$, $p > .05$, $\eta_p^2 = .03$. Additional analyses using t tests showed that that the adult group performed significantly better than each of the groups of children, both ps $<$ .05, but the groups of children did not differ significantly, $p > .05$. These analyses also showed that performance on the intervention no delay condition was significantly poorer than that in the other two conditions, and that performance on the time only condition was significantly better than that in the intervention plus time condition, all ps $< .02$. Further analyses examined whether participants chose the correct structure more often than would be expected by chance, given that the probability of choosing the correct picture was one third. Data from the two causal chain trials are analyzed together because performance was very similar on these trials; to compare performance on these trials against chance, we compared the number of participants getting 0, 1, or 2 causal chains correct against the number expected by chance using a goodness-of-fit χ^2 test.

For the intervention no delay condition, participants in all groups chose the correct structure for the common cause box more often than would be expected by chance, all binomial ps $< .001$. However, only the adult group chose the correct answer for the causal chain structures more often than would be expected by chance, $\chi^2(26) = 14.59$, $p < .001$. It can be seen from Figure 3a that on the causal chain boxes children tended to choose the common cause structure rather than the appropriate causal chain. Thus, children's above-chance performance on the common cause structure appears to be a result of a general preference to choose the common cause response in this condition.

In the time only condition, participants' performance was assessed in terms of whether they chose the causal structure that was suggested by the temporal pattern of the events (i.e., common cause for the synchronous schedule and causal chains for the sequential schedules; because there was no statistical information in this condition, there was no objectively correct answer). For all groups and all causal structures, participants chose the box suggested by the temporal cues more frequently than would be expected by chance, all ps $< .01$. For the intervention plus time condition, participants in all groups also chose the structure suggested by both the intervention cues and the temporal cues more often than chance, all ps $< .01$. In this condition, there was no consistent tendency to choose one of the

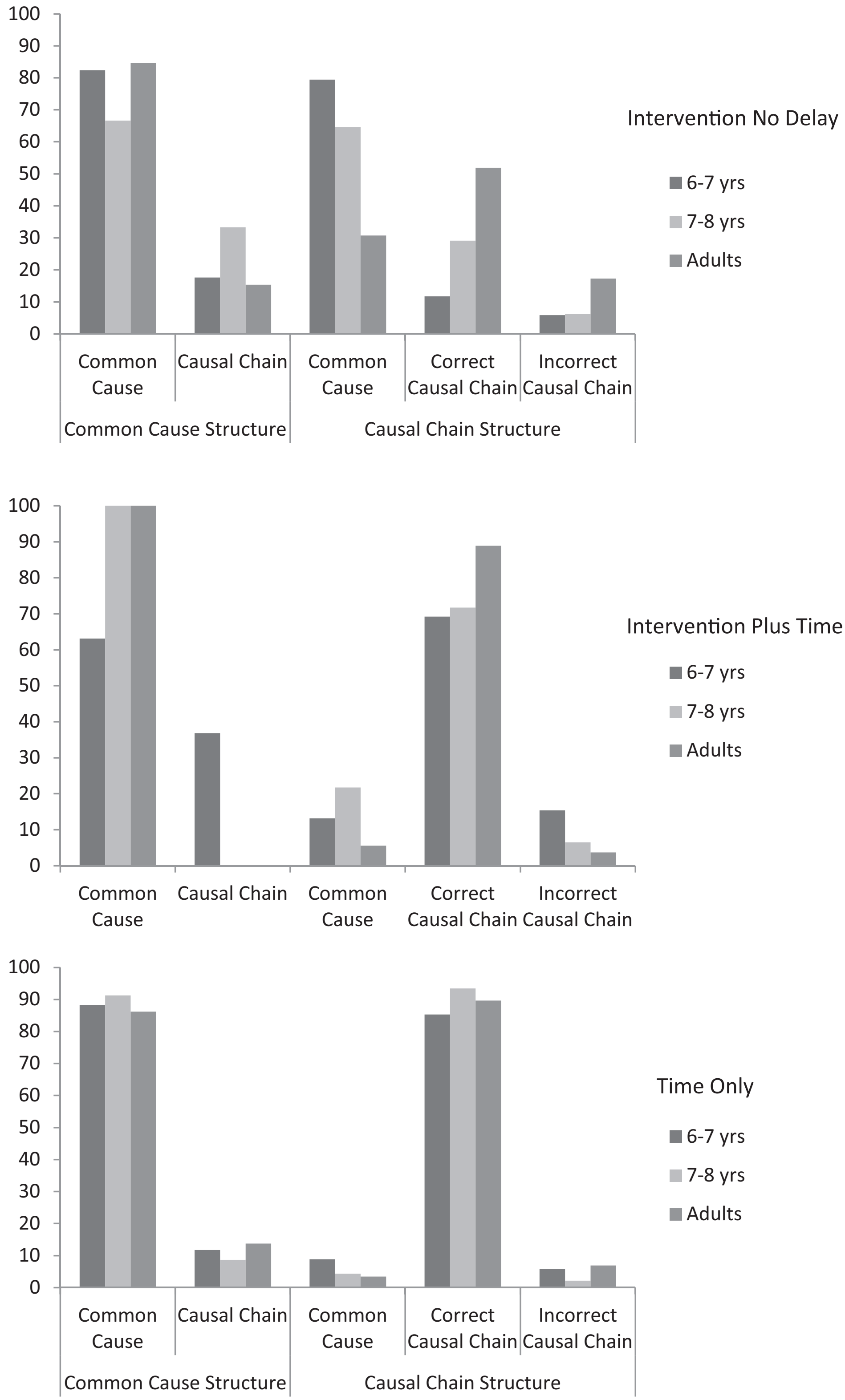

Figure 6. Percentages of causal model choices for each structure in Experiment 2 for each condition.

other structures when participants made an error on causal chain structures.

Discussion

Our findings suggest that children and adults readily use temporal pattern cues to infer a causal structure, and that this is true regardless of whether participants are provided with additional information from generative interventions. Note that in the time only condition, temporal pattern cues provided a basis for making causal structure judgments, but participants were not actually provided with sufficient information to make a veridical judgment about causal structure. It is possible for the structure to be a common cause even if the temporal pattern of events is sequential, that is, even if there is a delay between the operation of B and C. Conversely the synchronous temporal pattern of events is compatible with a causal chain structure: even if there is no delay between B and C it remains possible that B is a cause of C or vice versa. Nevertheless, as in previous studies (Burns & McCormack, 2009; Frosch et al., 2012), participants of all ages were willing to use temporal cues when no other information was available.

Children's performance was poor in the intervention no delay condition, despite the fact that the interventions provided statistical information in a way that would have allowed participants to learn whether or not each individual pair of nodes was causally linked. It might have been expected that even if participants had learned these pairwise links separately, they could have been able to discriminate between the three different options at test. For example, for an ABC chain, participants saw that moving B caused C to move but moving C did not cause B to move; only one of the three possible structures had an arrow from B to C. However, children were not above chance in correctly judging the structure of the causal chains, and although they gave correct responses to the common cause structure, this appeared to be due to a general tendency to give common cause responses (see Figure 6a) rather than due to children using conditional dependence information appropriately. One obvious explanation for participants' tendency to give common cause responses in the intervention no delay condition was that they again defaulted to using temporal pattern cues. In this condition, participants saw each component being operated by the experimenter, including A. When A was operated, B and C also occurred simultaneously. This may have suggested to participants that the structure was a common cause structure: although the temporal pattern was not identical to the synchronous structure used in the time only condition (there was no, or no perceptible, delay between the operation of A and that of B and C), nevertheless it may have been sufficiently similar to also suggest a common cause structure. In this sense, the findings of this experiment are highly consistent with the statistical information only condition in Experiment 1, in which all events happened simultaneously and children were also likely to judge that the causal chain structure was a common cause.

We note also that not only did children have difficulty using intervention information in the intervention only condition, the availability of such information also seemed to impair performance when combined with temporal information: Performance in the intervention plus time condition was somewhat worse than performance in the time only condition, at least for the child participants (see Figure 6). This may be because children were distracted by the experimenter's interventions on the system and found it difficult to combine intervention and temporal information together, even though temporal and intervention information were consistent with each other. This suggestion fits with the idea that children have difficulty integrating information from different sources, either across observations (as required in piecing together causal structure from pairwise causal links) or in this case different types of information from the same observations.

Experiment 3: Prevent-Then-Generate Interventions

The temporal schedule of the events in the intervention no delay condition of Experiment 2 was identical to that of Schulz et al. (2007) and Sobel and Sommerville (2009), who also used procedures in which the events occurred simultaneously. However, our study differed from theirs in that we did not show participants the consequences of preventative interventions—interventions that prevented a component in the system from operating. There are at least two reasons that people might find it easier to use information from preventative interventions. First, preventative interventions establish unambiguously whether or not the operation of one component is necessary for another component to operate. The system used in Experiment 2 was a closed, fully deterministic system, with events never occurring spontaneously, which means that, for example, if C occurred when B was operated by the experimenter, there had to be a causal link between B and C. Nevertheless, it is possible that participants either did not fully appreciate this, or that they believed that A was a common cause of B and C, but took this to be compatible with there being a separate causal link between B and C and thus selected the common cause picture. Showing participants preventative interventions can rule out the latter interpretation: For example, for an ABC causal chain, such an intervention can be used to show that C cannot work if B is disabled.

The second reason that using preventative interventions may lead to better performance is that it enabled us to provide participants with statistical information through interventions without additionally showing the temporal pattern of events that children seemed to have interpreted as evidence for a common cause structure in Experiments 1 and 2 (all three events A, B, and C happening simultaneously). In our third experiment, participants never saw A operating along with both B and C. Rather, the experimenter selectively disabled each component of the system in turn and demonstrated the effect that this had on the operation of the other components—what we have termed prevent-then-generate interventions. The result of this was that whenever A was operated, either B or C (but not both) had already been disabled by the experimenter. Removing the temporal information in Experiment 3 meant that participants only ever had one source of information with which to make their causal structure judgments, dependency information from observing interventions. We were interested in whether participants could make use of this information.

Participants

Eighty-eight children (42 girls) from three different school years took part in the experiment: 32 children aged 5- to 6-years-old (M = 72 months, *range* = 62–78 months), 29 children aged 6- to

7-years-old (M = 85 months, *range* = 80–90 months), and 27 children aged 7- to 8-years-old (M = 99 months, *range* = 94–104 months). Children were recruited from and tested in their schools. Twenty-five adults also participated, who were recruited and tested in the same manner as those in the previous experiment.

Materials

The same apparatus was used as in Experiment 1, with one additional prop. In order to disable components of the causal system, a metal bar with a miniature stop sign affixed to one end was used. This resembled a stop sign commonly used as a road traffic signal. The metal bar could be inserted into a small hole in any one of the components on the box's surface that could be lined up with a hole in the box itself to secure it, preventing it from rotating (see Figure 2; see also Frosch et al., 2012).

Procedure

Children were introduced to the box and pretrained on the meaning of the response pictures as in Experiment 1. Following this, participants were shown the stop sign, and the experimenter explained that it could be used to disable a component from moving by inserting it into a hole in a component. The experimenter demonstrated this by putting the stop sign into the hole in the A component on the first box and showing it made it physically impossible for A to move. The experimenter then asked children to tell her whether the box was on or off (it was off), and then switched the box on. Following this, she told children: "I'm going to stop this shape from moving (pointing to one of the shapes) and show you what happens to the other shapes. I will do that for all shapes. So, for each shape I am going to show you what happens to the other shapes when I stop it from moving." She then reminded children what their task was by saying "Remember, you've got to figure out which picture shows how this box really works. Are you ready now?" She then disabled one of the shapes (counterbalancing whether it was A, B, or C), and, for each of the remaining two shapes, demonstrated twice what happened when she operated the other shape, counterbalancing which other shape was operated first. She then disabled a different component and again for each of the other remaining shapes demonstrated twice what happened when she operated the other shape, and then repeated this for the third component in the system.

Thus, participants saw the effects of disabling A, B, and C on the operation of the other components. In three separate trials, children were shown the three different causal structures used in Experiment 2, and made their responses by selecting between the same set of three pictures. As in the first two experiments, a short unrelated distracter task was used in between demonstrating each structure, and children were told that each new box might work in the same way as the previous box or that it might work in a different way. The procedure for adults was very similar, although adults did not complete the comprehension pretraining.

Results and Discussion

Eleven children did not pass the comprehension questions the first time (five 5- to 6-year-olds, five 6- to 7-year-olds, and one 7- to 8-year-old), and were removed from the analyses although the pattern of findings remains unchanged if these children are included. Figure 7 shows the percentage of participants from each group who chose each response for the common cause and causal chain structures. An initial ANOVA on the number of correct responses (0–3) with a between-subjects factor of age group showed a main effect of age $F(3, 98) = 15.57$, $p < .001$, $\eta_p^2 = .32$; subsequent t tests showed that the adults performed significantly better than any of the groups of children (all ps $< .01$), and that the 7- to 8-year-olds performed marginally significantly better than the 5- to 6-year-olds (p = .06). Additional analyses were used to examine if each group of participants selected the correct response more often than would be expected by chance. The 5-to 6-year-old group did not select the correct response for the common cause structure more often than would be expected by chance, binomial $p > .05$; nor did they choose the correct structure for the causal chains more often than would be expected by chance $\chi^2(2) = 3.42$, $p > .05$. Neither the 6- to 7-year-olds nor the 7- to 8-year-olds selected the correct response to the common cause structure more often than would be expected by chance, both ps $> .05$, but both

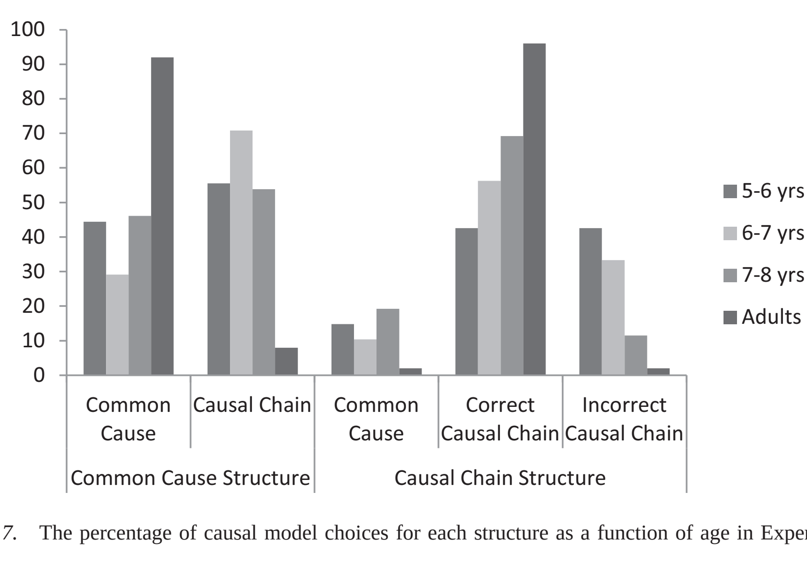

Figure 7. The percentage of causal model choices for each structure as a function of age in Experiment 3.

groups selected the correct responses for the causal chains more often than would be expected by chance, $\chi^2(2) = 11.72$, $p < .01$, and $\chi^2(2) = 48.77$, $p < .001$. The adult group were above chance on both the common cause structure, binomial $p < .001$, and the causal chain structures, $\chi^2(2) = 165.80$, $p < .001$.

The findings from this study contrast with those of Experiments 1 and 2, in that children were accurate at discriminating the causal chain structures, at least from 6- to 7 years, whereas in the statistical information only condition of Experiment 1 and the intervention no delay condition of Experiment 2 children performed accurately only on common cause structures. We have argued that the high levels of common cause responses in those conditions were due to the fact that children saw all three components operating simultaneously, a temporal pattern which they seem to interpret as indicating a common cause structure. Thus, children's accurate performance on common cause structures in Experiments 1 and 2 was not due to them making use of the statistical information provided to them. However, in this experiment, no such temporal cues were available to children. Nevertheless, the two older groups of children were able to use the information provided by the intervention demonstrations to give accurate responses in the causal chain trials. Note that this cannot be due to a response bias for a particular causal chain structure because to do well on causal chain trials participants had to give ABC responses for one chain and ACB responses for the other. Thus, the findings of this study, unlike those of Experiment 2, indicate that by around 6 years children can use intervention information in causal structure learning. The issue to be explained is why they did so only for the causal chain structures and not for the common cause structure; we return to this issue in the General Discussion.

General Discussion

Across three experiments, adults and children had to use different types of statistical information and/or temporal pattern information to discriminate between causal structures. The first aim of these studies was to examine how competent participants are at using statistical information in comparison to temporal pattern information. The findings of Experiments 1 and 2 consistently support the hypothesis that participants find it much easier to use temporal information. When temporal and statistical information were in conflict with each other, all groups of participants preferentially used temporal over the statistical information provided by observing the operation of a probabilistic system (Experiment 1). In Experiment 2, statistical information was provided by participants observing the experimenter making each node in the system activate in turn. Participants in all groups found it more difficult to use this statistical information than temporal pattern information. In Experiment 3, only statistical information was provided, by means of the experimenter demonstrating the effects of interventions that involved disabling one of the nodes in the system and making each of the other nodes activate in turn. Under these circumstances, adults and the older two groups of children were able to use the statistical information, at least for some trial types (causal chains).

The second aim of these studies was to examine whether there were age-related changes in participants' ability to use statistical information in causal structure learning in comparison with temporal pattern information. We predicted that children might find it particularly difficult to use statistical compared to temporal information. Our findings suggest that children do indeed find it difficult to use statistical information. Adults managed to use statistical information on its own to perform at above-chance levels in all three experiments. However, none of the groups of children did so in Experiments 1 and 2. Children appeared to rely primarily on temporal cues in those experiments. In Experiment 3, in which only statistical information was available, there was a developmental pattern: The two older groups of children were able to use intervention information to discriminate between causal chain structures. However, the younger children (5- to 6-year-olds) did not perform at above-chance levels on any trial type. Moreover, even the older children had difficulty with the common cause structure.

Integrating Information

In the introduction, we defined as an *information processing account* the claim that participants use temporal heuristics in their causal structure learning in preference over statistical information because of the processing demands of making inferences based on integrating information provided over a number of observations. We distinguished between this account and the idea that there might be an *inherent bias* to use temporal information. As noted in the introduction, these two explanations are not mutually exclusive, and it may be that both of them combine to explain our data. That is, children's difficulties in using statistical information in our tasks, and their preference for temporal information, may have two sources: They may still struggle with the information processing demands of integrating information across observations even when the tasks are simplified as in the current study, and this difficulty may be combined with an inherent bias toward use of temporal information.

However, the information processing account rather than the inherent bias account helps make sense of two other, potentially puzzling, findings from our study. Both of these findings can be interpreted as suggesting that information integration may prove challenging in causal structure learning, particularly for young children. The first is the somewhat surprising finding that participants, and again children in particular, performed worse in intervention plus time condition than the time only condition in Experiment 2. This finding is counterintuitive because exactly the same temporal cues were available in both conditions, and we have been suggesting that participants find it easy to use such cues. One interpretation of this finding, mentioned above, is that participants' performance is affected by an attempt to try to integrate statistical information from the generative interventions with the easy-to-use temporal information. Even attempting such integration of information may have lowered performance, especially in the groups of children.

The second puzzling finding comes from Experiment 3, in which the older two groups of children were able to use statistical information from prevent-then-generate interventions to make correct judgments about causal chain but not common cause structures. The fact that the older children can use statistical information to make causal chain judgments indicates that they can in principle use this type of information, which raises the issue of why they did not do so for the common cause structure. One possible explanation of this finding lies in the different requirements of these two structure types with regard to integrating information across observations. In Experiment 3, children could have given the correct answer for the causal chain structures by attending to only one demonstration type. For example, if the chain was an ABC chain, information provided on demonstrations in which B was disabled and A was operated would be sufficient to discriminate this structure from the other two available response options (the ACB causal chain and the common cause structure). However, this was not the case for the common cause structure, which required at least two demonstration types for answering correctly (e.g., showing that when B is disabled, C still moves when A is operated, and that when C is disabled B still moves when A is operated). Perhaps children's difficulties lay with piecing together information across observations, which was not essential for correct responses in the causal chain trials. Adults were at ceiling for both causal chain and common cause trials, indicating that they found this task trivially easy.

Additional studies that carefully examine the extent to which children can integrate information across observations and information types in causal structure learning are necessary to establish whether this is the correct explanation of performance. However, if these interpretations of our findings are correct, then they are consistent with the suggestion that processing limitations affect the ability to make use of statistical information. This suggestion also coheres with McCormack et al.'s (2013) recent findings that working memory abilities impact on children's causal learning. It may be worth extending this line of research to examine the relationships between working memory development and children's causal structure learning, because it seems likely that giving a complete account of causal structure learning (and its development) requires exploring the role of such processing factors.

Preferential Reliance on Temporal Information Over Statistical Information

So far, we have argued that children may be particularly likely to rely on temporal information because of their limited processing resources, although we have left open the possibility that they may also have an inherent bias for temporal information. The question remains, though, of how to explain the tendency of adults to also rely on temporal pattern information, despite the simplified nature of the tasks and adults' greater information processing resources. Although adults were able to use statistical information at above chance levels across all of the experiments, it is notable that many of them did not do so in Experiments 1 and 2. Adults, as well as children, used temporal pattern information in preference to statistical information in the incongruent condition in Experiment 1. Moreover, in the statistical information only condition of Experiment 1, and the intervention no delay condition of Experiment 2, only around half of the adult group used statistical information to infer causal chains. The other half seemed to use the temporal pattern derived from the observation that, when all components operated, they operated simultaneously, and infer a common cause. This was despite the fact that (e.g.) for an ABC causal chain in Experiment 1, they would have seen a number of trials in which A and B occurred without C and none in which A and C occurred without B, and in Experiment 2, they would have seen that operating B on its own made C activate.

One might argue that it is less plausible that adults failed to use the statistical information simply because of its information processing demands. After all, seeing that B makes C operate straightforwardly provides evidence for the existence of a causal link between B and C, and at test there was only one response option that includes such a link. The alternative possibility is that there is some inherent bias for temporal information over statistical information that makes it the case that adults as well as children can be reluctant to relinquish the inferences they have made on the basis of temporal information even in the light of contradictory statistical information. This raises the interesting question of what might underlie such an inherent bias and in what contexts we might expect to observe it.

As we pointed out in the introduction, White (2014) has argued that biases of this sort stem from the origins of causal cognition in human action. Rather than focusing on White's specific claims, we wish to broaden out the discussion so that it links with a generally accepted distinction between two different notions of causation itself. Woodward (2011) distinguishes between what he terms the geometrical-mechanical notion of causation and the difference-making notion of causation. This distinction actually has its origins in metaphysical debates about causation, but Woodward uses the terms to distinguish between two different ways psychologists commonly conceptualize what it is to represent a relation as causal. On the former notion of causation, causal relationships involve physical processes that connect causes to effects. This can be seen clearly in psychological accounts that place the idea of a mechanism connecting cause and effect at the heart of causal cognition; such accounts have featured in both the developmental literature and in the adult work on causal inference (for discussion, see Ahn, Kalish, Medin, & Gelman, 1995; Mayrhofer & Waldmann, 2014; Shultz, 1982; Shultz, Fisher, Pratt, & Rulf, 1986; Wolff, 2007). Other psychological research that draws on a geometrical-mechanical notion of causation is the Michottean tradition (Michotte, 1946/1963), which has examined the perceptual cues that lead to causal impressions when objects collide or interact. The idea here is in some circumstances our perceptual system uses simple spatiotemporal cues to track the physical processes that connect causes with effects.

On the second, difference-making notion of causation, a cause is something that makes a difference to whether or not an effect occurs. This notion links clearly to psychological accounts that emphasize the role of statistical information in causal inference, because whether a purported cause makes a difference to whether or not an effect occurs might be assessed by gathering information about the likelihood of the effect occurring in the presence and absence of the purported cause. In his own work, Woodward (2003) advocates a difference-making theory that defines causal relations as those in which unconfounded interventions on the cause lead to correlated changes in the effect. Various proponents of the causal Bayes net approach to causal structure learning explicitly align themselves to difference-making accounts (see contributions to Gopnik & Schulz, 2007, particularly Schulz, Kushnir, & Gopnik, 2007). As Woodward (2011) puts it, causal Bayes nets "are also naturally understood as difference-making in spirit, since such structures involve the representation of claims about how the probabilities of the values of effect variables will change depending upon changes of various sorts in the value of cause variables" (pp. 26–27).

Woodward (2011) argues that, at least in adult human causal cognition, both these modes of thinking about causation—the geometrical-mechanical and the difference-making notions—coexist. He suggests that typically they operate in tandem and are highly integrated with each other (see also Sloman & Lagnado's [2014] review of psychological work on causality for a similar argument). Such integration would be unsurprising because, in general, if there is a causal process connecting A and B, variations in the state of A will make a difference to the state of B. However, even if these two notions are well-integrated, it is possible that there are circumstances in which they could possibly lead to conflicting conclusions about causal relations, given a specific pattern of evidence. For example, there may be circumstances under which more weight is given to statistical evidence that indicates that manipulations of A make a difference to the state of B than to evidence that there is no causal process or mechanism connecting A and B, or there may be circumstances in which causal process information is weighted more heavily than such statistical evidence.

Empirical findings relevant to this issue come from two influential studies that have deliberately pitted statistical information against mechanism information. Shultz (1982) demonstrated that children will preferentially rely on their knowledge of a mechanism connecting two events over statistical information about covariation. Relatedly, Ahn, Kalish, Medin, and Gelman (1995) provided adult participants with sets of sentences describing events with two potential causal factors, and pitted the covariation information provided about one factor against the mechanism information provided about another factor. They found that participants placed considerably more weight on the mechanism information than the covariation information. Both these sets of findings can be interpreted as suggesting that the geometrical-mechanical way of thinking about causation may be more fundamental than the difference-making way (though see Danks, 2005; Newsome, 2003, for discussion). For present purposes, the key point is that Shultz's (1982) and Ahn et al.'s (1995) findings that participants give less weight to statistical information resonate with those from the current studies, although we compared use of statistical information to use of temporal information rather than information about mechanism.

Indeed, in our experiments, as in the vast majority of recent studies of causal inferences, participants were provided with no explicit information about the actual mechanisms that underpinned the relationships between the events. The mechanism that operated the apparatus was concealed from participants (a laptop computer that was hidden inside the box controlled each event). Nevertheless, though, it could be argued that a geometrical-mechanical account of causal thinking would predict that temporal information will be more salient than statistical information. This would be the case if temporal information is taken as an inherent feature of the causal processes connecting the events, resulting in participants placing considerable weight on such information. Thus, one interpretation of our findings is that participants—and children in particular—tend to use a geometrical-mechanical mode of thinking about causation, and that, at least in some circumstances, they use this mode of thinking over a difference-making mode of thinking about causation when the two are in conflict.

In our study, participants had no reason to attribute the temporal pattern of events to anything but the underlying causal mechanisms (cf. Lagnado & Sloman, 2006). In the absence of explicit mechanism information, it may be that participants treated temporal cues as a reflection of how the underlying mechanisms operated, and that this is why such cues were accorded more weight than statistical information. We do not know to what extent participants speculated about the underlying causal mechanisms in the current studies, but even if participants did not have any well-worked out hypotheses about the nature of these mechanisms, nevertheless they may have interpreted the temporal information as stemming from the underlying mechanisms and thus interpreted causal structure on the basis of the temporal pattern of the events. If this is correct, then the best interpretation of our findings is that the bias we observed for temporal pattern information stems from a general bias for information about causal process. One prediction of such an account is that the way in which participants rely on such a temporal cue will vary greatly depending on the nature of the mechanisms involved and participants' mechanism knowledge. Such findings would be consistent with Buehner's demonstrations that temporal contiguity is weighted differently as a cue to causation depending on what mechanism participants believe underlies the connection between a cause and its effect (e.g., Buehner & May, 2002, 2003, 2004; Buehner & McGregor, 2006).

Using Statistical Information: Is it Really so Difficult?

We have provided two possible, and not mutually exclusive, explanations of the finding that participants in our studies, and children in particular, seemed to rely on temporal cues and had difficulty using statistical information in our studies. One of these explanations focused on the information processing demands of using statistical information, and the other on the idea that participants may preferentially use information that they take to reflect the physical process connecting causes to effects. We want to finish by considering to what extent our findings are consistent with those of other studies that indicate that both adults and children may be adept at using statistical information in causal learning.

Our developmental findings might be taken to be particularly controversial given the large body of recent research on causal learning that has been conducted within the Bayesian tradition and has emphasized young children's competence in using statistical information (e.g., Gopnik, 2012; Gopnik et al., 2004; Gopnik & Wellman, 2012; Sobel, Tenenbaum, & Gopnik, 2004). We note that the Bayesian formal approach itself is a normative one, and theorists in this general tradition need not assume that children or indeed adults actually make inferences by processing statistical information. Rather, the claim is that performance will approximate to that of an idealized Bayesian learning system that updates hypotheses about causal structure using appropriate statistical information. However, Gopnik (2012) has argued that children do indeed extract statistical patterns from data in causal learning tasks, and that it is as a result of doing so that their performance "often resemble idealized Bayesian learners" (p. 1625).

The vast majority of studies of children's causal learning have used tasks in which participants need to make judgments about the causal efficacy of certain objects, rather than make causal structure judgments of the sort used in the current study. We are not arguing here that young children never use statistical information in causal learning (though see White, 2014, for a much more skeptical interpretation of developmental findings that have been interpreted as demonstrating children's use of such information). Rather, we interpret our findings as suggesting that it is not necessarily easy for them to do so. The claim that there are important developmental changes in children's ability to use statistical information has been around for a long time in the literature (e.g., Sedlak & Kurtz, 1981; Shaklee & Goldston, 1989; Shaklee, Holt, Elek, & Hall, 1988; Shultz & Ravinsky, 1977; Shultz, 1982; Siegler & Liebert, 1974). What our studies add to this earlier body of literature is the finding that (a) children can find it difficult to use statistical information in the context of causal structure learning, and (b) this is the case even in circumstances in which such information is provided as a result of very simple interventions on a deterministic system, as in Experiments 2 and 3. Explaining children's difficulties in using this statistical information in causal structure learning is likely to require properly understanding the demands that using such information places on children's limited cognitive resources.

Given our interpretation of children's difficulties, it may be that experimental manipulations that reduce the information processing demands of the task could strongly affect children's performance. We note that in the Sobel and Sommerville (2009) study mentioned in the introduction, unlike in the current study, children were carefully led through the process of constructing two competing hypotheses, reminded about these hypotheses, and then the experimenter showed children just one intervention and got children to focus on its outcome (whether A still caused C in the absence of B). We have argued in the introduction that children need not have properly grasped the causal structure of the three variables in the system to pass the task, because giving the correct answer did not require integrating pieces of information about pairwise causal links (the same argument applies to Sobel & Sommerville's [2010] study). Setting aside this argument, though, Sobel and Sommerville's (2009) procedure may have greatly reduced the information processing resources that children required to decide on whether two of the variables were causally linked. Such scaffolding may be necessary for young children to make use of even simple statistical information in causal structure learning.

Adults were at ceiling in using statistical information in Experiment 3, where that was the only source of information available. This finding is consistent with other studies showing that adults can indeed use statistical information to learn causal structure, particularly if the system is deterministic (Deverett & Kemp, 2011; Kushnir, Gopnik, Lucas, & Schulz, 2010). Put together with other findings in the literature, the findings of Experiment 3 suggest that when adults failed to use such information in Experiments 1 and 2, it was not because they were unable to do so. Rather, they preferred to use temporal pattern information. If this is correct, then it should be possible to demonstrate that adults can straightforwardly use the sort of statistical information provided in Experiments 1 and 2 if temporal pattern information was removed entirely (perhaps by only showing the outcomes of interventions, not the interventions happening in real time). Alternatively, it may be that participants can be encouraged through instruction to focus instead on statistical cues. Fernbach and Sloman (2009) have demonstrated that pretraining that allows participants extended practice with feedback in learning causal structure can affect the extent to which they default to relying on simpler heuristics. If adults' preference for temporal information is an inherent bias, rather than due to processing limitations, instructions or pretraining that encourage participants to shift their attention away from temporal cues might be expected to result in greater use of statistical information.

A further possibility is that the extent to which participants show an inherent bias for temporal pattern information, or causal process information more generally, will vary, even in the absence of instruction or pretraining, depending on the learning context in which they find themselves. In the majority of previous studies of causal structure learning (see Kushnir et al., 2010, for an exception), participants see the values of variables change on a computer screen. This means that there are no physical causal connections between the variables in question, and no causal process connecting these values (except extremely indirectly via the computer program running the task). Under these circumstances, it may be that participants are likely to more readily use statistical information. In the current task, there was an actual mechanical system and although the objects in the system were in fact controlled by a computer program, participants did not know this, and very simple mechanical connections could have underpinned the events that participants observed. It may be that when the causal systems being observed are real rather than just depicted on a computer screen, participants are more likely to weigh temporal pattern/causal process information more heavily than statistical information.

Computational Approaches to Causal Induction From Temporal Order Information

A major finding from the current studies is that people use temporal pattern information even when it appears to conflict with statistical information and even when the learning task is simplified. This seems, at least at first sight, inconsistent with the claim that people's causal induction conforms to optimal Bayesian learning. However, how well people's inferences fit with a Bayesian model cannot be properly assessed without a broader computational framework that includes both: (a) a model for how causality should be inferred from temporal information, and (b) a model for integrating temporal and statistical information. Although Bayesian computational models for learning from statistical data are relatively well worked out, and theorists acknowledge the important role of temporal information in causal induction (Griffiths & Tenenbaum, 2009; Holyoak & Cheng, 2011), there is currently no established computational framework for inferring causality from temporal order, nor for combining temporal and statistical information.

The findings in the current article present an important challenge to computational models, but also an opportunity to extend successful models of learning from statistical data to include a broader range of information such as temporal order and delays. One step in this direction is provided by Pacer and Griffiths (2012), who outline criteria for a rational framework of causal induction from continuous time information (e.g., delays and rates). However, in its current formulation their framework does not deal with temporal order per se, nor does it directly apply to the three variable structures used in our current experiments.

Bramley, Gerstenberg, and Lagnado (2014) present a modeling framework that focuses directly on the use of temporal order and delays for inferring three-variable causal structures. In their experiments people infer structures from patterns of temporal information where the statistical information is held constant (i.e., participants view clips where variables are always activated, but the time order and delays between events are systematically varied). They explore Bayesian and bounded Bayesian models and show that the human data is best fit by a model which assumes that causes precede their effects, and do not occur simultaneously with them (e.g., if A and B occur at the same time, then it is unlikely that A causes B or B causes A). Thus simultaneously occurring events are most likely to be effects of an earlier common cause, or joint causes of a later effect. They also found that people updated their beliefs about causal structures more conservatively than predicted by a Bayesian model.

The assumption that causes precede their effects fits with the results reported in the current article. Thus, in the incongruent condition in Experiment 1, when people see A occur, followed by the simultaneous co-occurrence of B and C, people choose the common cause structure even when presented with statistical data from a chain structure. The fact that B and C occur at the same time seems to rule out the possibility that B causes C or vice versa. Moreover, even in the statistical information only condition in Experiment 1, the assumption might still be operative. A is already marked out as a cause (all response options have A as the root cause), and B and C activate simultaneously, making the common cause response option more consistent with the no-simultaneity assumption than the chain. Similar arguments apply to the findings in Experiment 2. Finally, in Experiment 3 there was no temporal pattern information and the three variables were never shown occurring together. Here the no-simultaneity assumption cannot be used to rule out links between B and C, and indeed the common cause structure was no longer preferentially endorsed.

Overall, the inferences drawn from the temporal patterns in our experiments are consistent with the simple assumption that causes precede, and thus are unlikely to occur at exactly the same time as, their effects. This assumption is, however, defeasible—when all three variables occur simultaneously, and A is marked out as the root cause (by the response options or by an actual intervention), people tend to choose a common cause model, allowing that A can cause B (and C) at the same time. Further work is needed to develop computational approaches in more detail, and allow for the integration of both temporal and statistical information. This will also serve as a guideline for developing more appropriate psychological models.

Summary and Conclusions

We have demonstrated that participants, and children in particular, prefer or find it easier to use temporal pattern information over statistical information in causal structure learning, and that this is the case even when statistical information is provided in an accessible way that facilitates structurally local computations. There are clear developmental improvements in the ability to use statistical information; this is the case even if statistical information is not in competition with temporal information (Experiment 3), indicating a need to pinpoint the cognitive limitations that affect young children's causal structure learning. We have interpreted our findings with adults as suggesting that there may be an inherent bias toward use of temporal pattern information that may stem from a more general preference for information about the process connecting causes with their effects over statistical information. However, such a bias is as yet not fully understood.

References

Ahn, W. K., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition, 54,* 299–352. http://dx.doi.org/10.1016/0010-0277(94)00640-7

Bramley, N. R., Gerstenberg, T., & Lagnado, D. A. (2014). The order of things: Inferring causal structure from temporal patterns. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 236–241). Austin, TX: Cognitive Science Society.

Buehner, M. J., & May, J. (2002). Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking & Reasoning, 8,* 269–295. http://dx.doi.org/10.1080/13546780244000060

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56,* 865–890. http://dx.doi.org/10.1080/02724980244000675

Buehner, M. J., & May, J. (2004). Abolishing the effect of reinforcement delay on human causal learning. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 57,* 179–191. http://dx.doi.org/10.1080/02724990344000123

Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking & Reasoning, 12,* 353–378. http://dx.doi.org/10.1080/13546780500368965

Burns, P., & McCormack, T. (2009). Temporal information and children's and adults' causal inferences. *Thinking & Reasoning, 15,* 167–196. http://dx.doi.org/10.1080/13546780902743609

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

Cooper, G. F. (1999). An overview of the representation and discovery of causal relationships using Bayesian networks. In C. Glymour & G. F. Cooper (Eds.), *Computation, causation, and discovery* (pp. 4–62). Menlo Park, CA: AAAI Press and MIT Press.

Danks, D. (2005). The supposed competition between theories of human causal inference. *Philosophical Psychology, 18,* 259–272. http://dx.doi.org/10.1080/09515080500169371

Deverett, B., & Kemp, C. (2011). Learning deterministic causal networks from observational data. *Proceedings of the 34th Annual Conference of the Cognitive Science Society,* 288–293.

Fernbach, P. M., & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 678–693. http://dx.doi.org/10.1037/a0014928

Frosch, C. A., McCormack, T., Lagnado, D. A., & Burns, P. (2012). Are causal structure and intervention judgments inextricably linked? A developmental study. *Cognitive Science, 36,* 261–285. http://dx.doi.org/10.1111/j.1551-6709.2011.01208.x

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669. http://dx.doi.org/10.1037/0033-295X.103.4.650

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science, 337,* 1623–1627. http://dx.doi.org/10.1126/science.1223416

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111,* 3–32. http://dx.doi.org/10.1037/0033-295X.111.1.3

Gopnik, A., & Schulz, L. (2007). *Causal learning: Psychology, philosophy, and computation.* Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195176803.001.0001

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138,* 1085–1108. http://dx.doi.org/10.1037/a0028044

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review, 116,* 661–716. http://dx.doi.org/10.1037/a0017201

Hagmayer, Y., Sloman, S., Lagnado, D. A., & Waldman, M. R. (2007). Causal reasoning through intervention. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 86–100). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195176803.003.0007

Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology, 16,* 65–93. http://dx.doi.org/10.1016/0010-0285(84)90004-5

Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.

Halford, G. S., & Andrews, G. (2006). Reasoning and problem solving. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology: Vol. 2. Cognitive, language and perceptual development* (6th ed., pp. 557–608). Hoboken, NJ: Wiley.

Heckerman, D. (1998). A tutorial on learning with Bayesian networks. In M. I. Jordan (Ed.), *Learning in graphical models* (pp. 301–354). Cambridge, MA: MIT Press. http://dx.doi.org/10.1007/978-94-011-5014-9_11

Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology, 62,* 135–163. http://dx.doi.org/10.1146/annurev.psych.121208.131634

Kushnir, T., Gopnik, A., Lucas, C., & Schulz, L. (2010). Inferring hidden causal structure. *Cognitive Science, 34,* 148–160. http://dx.doi.org/10.1111/j.1551-6709.2009.01072.x

Lagnado, D. A., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 856–876. http://dx.doi.org/10.1037/0278-7393.30.4.856

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 451–460. http://dx.doi.org/10.1037/0278-7393.32.3.451

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

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review, 115,* 955–984. http://dx.doi.org/10.1037/a0013256

Mayrhofer, R., & Waldmann, M. R. (2011). Heuristics in covariation-based induction of causal models: Sufficiency and necessity priors. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the Thirty-Third Annual Conference of the Cognitive Science Society* (pp. 3110-3115). Austin, TX: Cognitive Science Society.

Mayrhofer, R., & Waldmann, M. R. (2014). Agents and causes: Dispositional intuitions as a guide to causal structure. *Cognitive Science*. Advance online publication. http://dx.doi.org/10.1111/cogs.12132

McCormack, T., Butterfill, S., Hoerl, C., & Burns, P. (2009). Cue competition effects and young children's causal and counterfactual inferences. *Developmental Psychology, 45,* 1563–1575. http://dx.doi.org/10.1037/a0017408

McCormack, T., Simms, V., McGourty, J., & Beckers, T. (2013). Blocking in children's causal learning depends on working memory and reasoning abilities. *Journal of Experimental Child Psychology, 115,* 562–569. http://dx.doi.org/10.1016/j.jecp.2012.11.016

Meder, B., Gerstenberg, T., Hagmayer, Y., & Waldmann, M. R. (2010). Observing and intervening: Rational and heuristic models of causal decision making. *The Open Psychology Journal, 3,* 119–135.

Michotte, A. (1946/1963). *The perception of causality*. London, UK: Methuen & Co Ltd.

Newsome, G. L. (2003). The debate between current versions of covariation and mechanism approaches to causal inference. *Philosophical Psychology, 16,* 87–107. http://dx.doi.org/10.1080/0951508032000067725

Pacer, M., & Griffiths, T. L. (2012). Elements of a rational framework for continuous-time causal induction. *Proceedings of the 34th Annual Conference of the Cognitive Science Society*.

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

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology, 64,* 93–125. http://dx.doi.org/10.1016/j.cogpsych.2011.10.003

Rottman, B. M., Kominsky, J. F., & Keil, F. C. (2014). Children use temporal cues to learn causal directionality. *Cognitive Science, 38,* 489–513. http://dx.doi.org/10.1111/cogs.12070

Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science, 10,* 322–332. http://dx.doi.org/10.1111/j.1467-7687.2007.00587.x

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

Schulz, L. E., & Sommerville, J. (2006). God does not play dice: Causal determinism and preschoolers' causal inferences. *Child Development, 77,* 427–442. http://dx.doi.org/10.1111/j.1467-8624.2006.00880.x

Sedlak, A. J., & Kurtz, S. T. (1981). A review of children's use of causal inference principles. *Child Development, 52,* 759–784. http://dx.doi.org/10.2307/1129077

Shaklee, H., & Goldston, D. (1989). Development in causal reasoning—Information sampling and judgment rule. *Cognitive Development, 4,* 269–281. http://dx.doi.org/10.1016/0885-2014(89)90009-9

Shaklee, H., Holt, P., Elek, S., & Hall, L. (1988). Covariation judgment - improving rule use among children, adolescents, and adults. *Child Development, 59,* 755–768. http://dx.doi.org/10.2307/1130574

Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology, 61,* 273–301. http://dx.doi.org/10.1146/annurev.psych.093008.100519

Shultz, T. R. (1982). Rules for causal attribution. *Monographs of the Society for Research in Child Development, 47* (1, Serial No.194).

Shultz, T. R., Fisher, G. W., Pratt, C. C., & Rulf, S. (1986). Selection of causal rules. *Child Development, 57,* 143–152. http://dx.doi.org/10.2307/1130646

Shultz, T. R., & Ravinsky, F. B. (1977). Similarity as a principle of causal inference. *Child Development, 48,* 1552–1558. http://dx.doi.org/10.2307/1128518

Siegler, R. S., & Liebert, R. M. (1974). Effects of contiguity, regularity, and age on children's causal inferences. *Developmental Psychology, 10,* 574–579. http://dx.doi.org/10.1037/h0036594

Simms, V., McCormack, T., & Beckers, T. (2012). Additivity pretraining and cue competition effects: Developmental evidence for a reasoning-based account of causal learning. *Journal of Experimental Psychology: Animal Behavior Processes, 38,* 180–190. http://dx.doi.org/10.1037/a0027202

Sloman, S. A., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science, 29,* 5–39. http://dx.doi.org/10.1207/s15516709cog2901_2

Sloman, S. A., & Lagnado, D. (2014). Causality in thought. *Annual Review of Psychology*. Advance online publication.

Sobel, D. M., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition, 34,* 411–419. http://dx.doi.org/10.3758/BF03193418

Sobel, D. M., & Sommerville, J. A. (2009). Rationales in children's causal learning from others' actions. *Cognitive Development, 24,* 70–79. http://dx.doi.org/10.1016/j.cogdev.2008.08.003

Sobel, D. M., & Sommerville, J. A. (2010). The importance of discovery in children's causal learning from interventions. *Frontiers in Psychology, 1,* 176. http://dx.doi.org/10.3389/fpsyg.2010.00176

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28,* 303–333.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation prediction and search*. New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-1-4612-2748-9

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27,* 453–489. http://dx.doi.org/10.1207/s15516709cog2703_6

Taylor, E. G., & Ahn, W. K. (2012). Causal imprinting in causal structure learning. *Cognitive Psychology, 65,* 381–413. http://dx.doi.org/10.1016/j.cogpsych.2012.07.001

Waldmann, M. R., Cheng, P. W., Hagmayer, Y., & Blaisdell, A. P. (2008). Causal learning in rats and humans: A minimal rational model. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind. Prospects for Bayesian cognitive science* (pp. 453–484). Oxford, UK: University Press. http://dx.doi.org/10.1093/acprof:oso/9780199216093.003.0020

Waldmann, M. R., Hagmayer, Y., & Blaisdell, A. P. (2006). Beyond the information given: Causal models in learning and reasoning. *Current Directions in Psychological Science, 15,* 307–311. http://dx.doi.org/10.1111/j.1467-8721.2006.00458.x

White, P. A. (2006). How well is causal structure inferred from cooccurrence information? *The European Journal of Cognitive Psychology, 18,* 454–480. http://dx.doi.org/10.1080/09541440500264861

White, P. A. (2014). Singular clues to causality and their use in human causal judgment. *Cognitive Science, 38,* 38–75. http://dx.doi.org/10.1111/cogs.12075

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General, 136,* 82–111. http://dx.doi.org/10.1037/0096-3445.136.1.82

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford, UK: Oxford University Press.

Woodward, J. (2011). A philosopher looks at tool use and causal understanding. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 18–50). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199571154.003.0002

Received August 5, 2013
Revision received June 2, 2014
Accepted June 4, 2014 ■