Developmental Psychology
2009, Vol. 45, No. 6, 1563–1575

0012-1649/09/$12.00 DOI: 10.1037/a0017408

Cue Competition Effects and Young Children's Causal and Counterfactual Inferences

Teresa McCormack
Queen's University Belfast

Stephen Butterfill and Christoph Hoerl
University of Warwick

Patrick Burns
University of Birmingham

The authors examined cue competition effects in young children using the blicket detector paradigm, in which objects are placed either singly or in pairs on a novel machine and children must judge which objects have the causal power to make the machine work. Cue competition effects were found in a 5- to 6-year-old group but not in a 4-year-old group. Equivalent levels of forward and backward blocking were found in the former group. Children's counterfactual judgments were subsequently examined by asking whether or not the machine would have gone off in the absence of 1 of 2 objects that had been placed on it as a pair. Cue competition effects were demonstrated only in 5- to 6-year-olds using this mode of assessing causal reasoning.

Keywords: causal learning, counterfactual reasoning, blocking, cue competition effects

Imagine your task is to learn which foodstuffs cause an allergy in a patient, a task commonly used in studies of adult causal cognition (e.g., Aitken, Larkin, & Dickinson, 2001; Dickinson, 2001; Le Pelley & McLaren, 2003). You find out that the patient has an allergic reaction after he (or she) has eaten cheese. Can you assume that cheese caused the reaction? Or, to put things counterfactually, would the patient still have had the reaction if he hadn't eaten cheese? How these questions should be answered depends on whether the patient has eaten other foods along with cheese and whether the allergic reaction has occurred in circumstances when cheese has not been eaten. In other words, it depends on what other cues might be competing in explaining why the allergy has occurred. For example, assume that the patient eats some nuts, gets the allergic reaction, and then eats some cheese and nuts together and also gets the allergic reaction. Under such circumstances, initially finding out that the patient has the reaction after eating nuts is likely to block new learning that cheese causes the reaction, because the cheese is eaten along with a foodstuff that is already known to cause the allergy. This type of effect is often referred to as blocking (e.g., Chapman & Robbins, 1990; Kruschke & Blair, 2000; Le Pelley, Oakeshott, & McLaren, 2005; Shanks, 1985) and is one of family of phenomena known as cue competition effects (De Houwer, Beckers, & Vandorpe, 2005; Shanks, 2007). Studies examining such phenomena demonstrate that causal judgments are not determined solely by how frequently a cue has occurred in the presence of outcome. Rather, the status of other cues is taken into account. The related phenomenon of causal discounting has been examined extensively by social psychologists (Kelley, 1972; Kruglanski, Schwartz, Maides, & Hamel, 1978). This research has explored whether adults take into account the role of other potential causes when attributing a motivation to behavior and has demonstrated that the role of a putative motivation is discounted if another plausible cause is also present. For example, if we already know that someone is very busy and preoccupied, we may discount the possibility that he is ignoring us because we have offended him.

Some relevant developmental research has already been been conducted on these issues. There is a long-standing research tradition on the development of causal discounting, with some research suggesting that if a potential alternative cause is made very salient, even preschoolers appear to discount certain motivations (Aloise & Miller, 1991; Miller & Aloise, 1989). Moreover, recent characterizations of causal learning in young children have also emphasized that their causal judgments are not a result of a simple association between a cue and an outcome based just on how frequently the cue and outcome co-occur, because children show sensitivity to the extent to which the outcome occurs in the absence of that cue and the relation of other cues to the outcome (Gopnik et al., 2004; Gopnik & Schulz, 2004; Gopnik, Sobel, Schulz, & Glymour, 2001; Schulz, Bonawitz, & Griffiths, 2007; Sobel, Tenenbaum, & Gopnik, 2004). However, there are still considerable gaps in our knowledge of whether young children's learning

Teresa McCormack, School of Psychology, Queen's University Belfast, Belfast, Northern Ireland; Stephen Butterfill and Christoph Hoerl, Department of Philosophy, University of Warwick, Coventry, United Kingdom; Patrick Burns, School of Psychology, University of Birmingham, Edgbaston, Birmingham, United Kingdom.

This research was supported by an interdisciplinary grant on causal understanding from the Arts and Humanities Research Council, UK, and a grant from the Experimental Psychology Society. We are grateful to Michelle Rankin, Helena McCormack, and Olive Taylor for assistance with data collection; to Fred Maddalena for constructing the equipment; and to Colin Sparrow for advice on Bayesian calculation.

Correspondence concerning this article should be directed to Teresa McCormack, School of Psychology, Queen's University Belfast, David Keir Building, 18-30 Malone Road, Belfast BT9 5BN, Northern Ireland. E-mail: t.mccormack@qub.ac.uk

embodies cue competition effects and whether they change developmentally (though see Beckers, Van den Broeck, et al., 2005; Beckers, Vandorpe, Debeys, & De Houwer, 2009; Sobel & Kirkham 2006; Sobel et al., 2004). Our primary aim in the studies reported in this paper was to conduct a more systematic developmental study of certain types of cue competition effects than has previously been reported.

Studies With Adults of Cue Competition Effects

Here we highlight two important manipulations that have been carried out in studies of cue competition effects in adults that have not yet been extensively studied in children; both of these manipulations were examined in our studies. First, one of the basic factors that studies of causal learning with adults have varied is whether or not a competing cue is itself paired with the effect. As we have described, in blocking procedures, the competing cue is independently paired with the effect (e.g., in our introductory example, an allergy occurs when nuts are eaten on their own; these will be labeled A+ demonstrations), and then a compound of two cues is paired with the effect (nuts and cheese in our example; labeled AB+ demonstrations). The initial A+ demonstrations are likely to block subsequent learning of a relation between B (cheese) and the allergy. However, in other procedures the competing cue is shown not to be paired with the effect (e.g., the hypothetical patient is given nuts and no allergy occurs; labeled A− demonstrations) before the compound cue AB+ is presented. Under such circumstances, participants are particularly likely to judge the target cue B (cheese) as causal; for present purposes we will refer to this as a generative effect of the initial A− demonstrations on learning (such effects are frequently referred to as prevention from overshadowing in the adult causal learning literature, but we will adopt this simpler terminology). Normatively, this type of judgment makes sense because once A is known not to be causal, B has to have caused the outcome (at least in deterministic scenarios). Table 1 illustrates the key trial structures for each of these cue competition effects. Note that when demonstrating both of these cue competition effects, one should compare responses to B to those from control trials involving two other novel cues (e.g., bread and strawberries in a food allergy task) in which neither cue is independently paired with the effect, in order to show that the individual pairings (A+ and A− demonstrations) have affected the causal ratings for B (De Houwer et al., 2005).

The second manipulation used extensively in studies of cue competition effects with adults has been to vary the order in which participants receive information about cues (e.g., Chapman, 1991; Kruschke & Blair, 2000; Lovibond, Been, Mitchell, Bouton, & Frohardt, 2003). As described in our example, in what we will henceforth label a forward blocking procedure, participants see a cue A (e.g., nuts) paired with an effect in advance of a compound of two cues AB (e.g., nuts and cheese) paired with the effect. By contrast, in a backward blocking procedure the phase involving AB+ presentations comes first, followed by a second phase in which A+ presentations are given (see Table 1). Backward blocking is observed under at least some conditions (e.g., Beckers, De Hower, Pineño, & Miller, 2005; Kruschke & Blair, 2000; Lovibond et al., 2003), as are what we will call backward generative effects, when participants see A− demonstrations after AB+ demonstrations (e.g., Beckers, De Houwer, et al., 2005; Chapman, 1991; this type of effect is often referred to as release from overshadowing).

Table 1
Presentations of Stimuli Used to Demonstrate Various Cue Competition Effects

Effect	First learning phase	Second learning phase
Forward blocking	A+	AB+
Forward generative	A−	AB+
Backward blocking	AB+	A+
Backward generative	AB+	A−
Control trials	CD+ (backward designs)	CD+ (forward designs)

Note. This table does not show the full set of stimuli used in any given experiment.

Such backward learning effects have been accorded much theoretical significance, because they rule out some simple but intuitively plausible models of causal learning. Forward blocking is readily explained by associative models of causal learning that assume that there is a change in the strength of an association between a cue and an effect only if the occurrence of the effect is not already predicted on the basis of the cues present (Dickinson, 2001; Rescorla & Wagner, 1972; such models also readily predict forward generative learning). For example, in forward blocking during AB+ demonstrations the effect is already predicted by the presence of cue A, which was previously shown paired with the effect; thus, such models assume that new learning of an association between B and the outcome is blocked. However, basic associative models do not predict backward blocking: According to basic associative models of causal learning, no learning about B should take place during the second phase because B is not present. In fact there has been considerable debate over how backward blocking should be modeled (e.g., De Houwer & Beckers, 2002; Denniston, Savastano, & Miller, 2001; Dickinson & Burke, 1996; Kruschke, 2006; Van Hamme & Wasserman, 1994).

On one approach, attempts have been made to develop models that retain some of the basic principles of associative accounts but make additional assumptions that allow them to explain such findings (Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994). An alternative, and now popular, approach has been to reject associative accounts and develop quite different ways of explaining causal learning. What are known as higher order theories of causal learning assume that processes similar to those that are normally considered reasoning processes, rather than learning depending on the mere formation of associations between cues and outcomes, are employed in such tasks (for review and discussion, see De Houwer, 2009; De Houwer et al., 2005; Pineño & Miller, 2007; Shanks, 2007).

Note that a number of quite different theories, which employ differing levels of description, could potentially be categorized as higher order theories (e.g., those of De Houwer et al., 2005; Gopnik et al., 2004; Waldmann, Hagmayer, & Blaisdell, 2006; for a discussion, see Pineño & Miller, 2007). Of particular interest from a developmental perspective are approaches that view cue competition effects on standard causal learning tasks as a result of processes that are controlled, effortful, and akin to inferential reasoning processes (De Houwer, 2009; De Houwer et al., 2005). Studies have shown that cue competition effects in adults are selectively disrupted if a secondary task has to be completed during learning, suggesting that such effects rely on working memory (e.g., De Houwer & Beckers, 2003; Vandorpe, De Houwer, & Beckers, 2005), are highly susceptible to task instructions regarding the nature of the cues (e.g., Waldmann, 2000, 2001), and are modulated by factors that might be expected to influence deductive reasoning (e.g., Beckers, De Houwer, et al., 2005; Lovibond et al., 2003). This range of findings is potentially important, because it naturally leads to the prediction that we might expect developmental differences in the extent to which cue competition effects are observed. Alternatively, if one managed to demonstrate such effects in very young children, this might call into question the extent to which such effortful reasoning processes necessarily underpin cue competition effects in humans.

Developmental Studies of Cue Competition Effects

Given the role of cue competition effects in debates over the nature of causal learning and the potential for developmental findings to contribute to such debates, it is important to review what we currently know about the nature of such effects in children's learning. Beckers, Van den Broeck, et al.'s (2005) study examined forward and backward blocking in 4- and 8-year-old children. In a single experiment they found blocking at modest though statistically significant levels and did not appear to find age effects; nor did they find an effect of trial order (i.e., levels of forward and backward blocking were similar). Sobel et al. (2004) examined backward blocking more extensively in 3- and 4-year-olds using the blicket detector paradigm. In this paradigm, cues are objects that either set off the detector (i.e., are blickets) or do not set off the detector (i.e., are not blickets). Versions of this paradigm have been used extensively in recent studies of children's causal learning, and indeed we use this paradigm in the studies reported here. Sobel et al. reported backward blocking in both age-groups, although blocking was more marked in their 4-year-old group. Moreover, Sobel and Kirkham (2006; Experiment 1) found some evidence of a generative effect, although not backward blocking, in the causal learning of children as young as 2 years of age when a simple forced-choice measure was used. Note, however, that the primary aim in Sobel et al. (2004) and Sobel and Kirkham (2006) was not to systematically examine whether children show backward learning per se but rather to examine whether children were sensitive to conditional probability information. That is, the researchers were interested, for example, in whether children were sensitive to whether or not a potential cause was related only to the effect conditional on the presence of another cause (see Sobel & Kirkham, 2007). Although the findings of Sobel et al. suggest that young children are indeed sensitive to such information (see also Kushnir & Gopnik, 2005, 2007), questions remain about the nature of the effects demonstrated in their studies. This is because they directly compared blocking trials to generative trials rather than compared backward blocking trials to the type of control trials shown in Table 1. As has been discussed, in studies of adult causal learning, judgments of the causal strength of B in generative trials are typically high, higher than, for example, those given to a paired cue from control trials. Thus, the fact that children were more likely to judge that B was a blicket in generative trials than in backward blocking trials does not allow us to be confident that blocking was observed.

Beckers et al. (2009) addressed this issue by examining 3-year-olds' performance on a blicket detector task in which judgments for backward generative and blocking cues were compared with those given to cues from control trials. The results were mixed: Beckers et al. found no evidence for either backward blocking or generative learning in this age-group when participants were asked to judge whether or not objects were blickets. However, they did find some evidence for generative learning but not backward blocking when using a forced-choice measure in which children were asked to choose an object to make the detector activate. In summary, although some evidence suggests that cue competition effects can be observed in young children, the findings from the small number of published studies do not allow us to be confident that this is the case and have not always been consistent. Moreover, it is not clear whether cue competition effects involving backward designs show the same developmental profile as do cue competition effects using forward procedures. We designed the present series of experiments to examine these issues.

Counterfactual and Causal Reasoning

In describing blocking in our introductory example, we raised the possibility of asking participants a counterfactual rather than a causal question: Rather than asking whether or not cheese caused the allergic reaction, we could ask whether the allergic reaction would have happened if the patient hadn't eaten cheese. There is a long history of history of debate in both philosophy and psychology regarding the relation between causal and counterfactual thinking (Collins, Hall, & Paul, 2004; Hagmayer, Sloman, Lagnado, & Waldmann, 2007; Hart & Honoré, 1959/1985; Lewis, 1973; Mackie, 1974; Mandel, 2003; Woodward, 2003, 2007), with some developmentalists arguing either that counterfactual reasoning underpins causal judgments (Harris, German, & Mills, 1996) or, more recently, that the sorts of representations that underpin causal judgments should also support counterfactual ones (Gopnik & Schulz, 2007; see also Woodward, 2007). In studies of the sort of cue competition effects we have described, the trial structure of the tasks means that participants have to make a causal judgment about a cue B that they have never seen paired on its own with or without the outcome (that is, they have seen B only in the presence of another cue A). In such tasks, there is a counterfactual scenario that would provide a direct information about B's causal status, one in which B occurs without A, but this scenario has not been observed by participants. The implication of this is that there is a counterfactual question—effectively, would the outcome have occurred in the absence of A, in a novel counterfactual scenario in which only B was present—that closely parallels the causal question. In fact, some psychologists who subscribe to higher order theories of causal reasoning have argued that there might be a particularly close relation between counterfactual and causal judgments in cue competition tasks, with Mitchell, Lovibond, and Condoleon (2005) suggesting that in blocking procedures participants consider unobserved counterfactual scenarios in coming to their conclusion about the causal status of cues.

Although we do not want to argue here that counterfactual reasoning underpins cue competition effects, such as blocking, the issue of whether children could answer counterfactual questions that have parallels to causal questions is important both in the light of recent theoretical accounts of causal representation that stress commonalities in the representations underpinning causal and counterfactual judgments (Gopnik & Schulz, 2007; Hagmayer et al., 2007; Sloman & Lagnado, 2005) and given ongoing debates about whether young children are capable of genuinely counterfactual thought (e.g., Beck, Robinson, Carroll, & Apperly, 2006; German, 1999; German & Nichols, 2003; Perner, Sprung, & Steinkogler, 2004). Experiments 3 and 4 reported below contribute to this debate because they assess young children's counterfactual reasoning about never-observed scenarios in a novel context. This design rules out the possibility that children could respond simply on the basis of general knowledge rather than counterfactual reasoning. For example, in Harris et al.'s study (1996) children were asked counterfactual questions about what would have happened if a doll had taken off her muddy shoes before walking across a floor. It is possible that children could have answered such questions correctly by simply drawing on their general knowledge that floors stay clean if you take off your shoes, rather than thinking counterfactually about the experimental scenarios actually observed. As Perner et al. (2004) put it, children could derive the answers to such questions "without recourse to the actual sequence of events" (p. 198). The questions used in our studies could not have been answered in this way and by necessity involved reasoning about unfamiliar scenarios and taking into account what had just happened.

We also note that asking counterfactual questions about never-observed scenarios can be seen as the linguistic analogue of eliciting certain action-based responses in a blicket detection task. In some previous developmental experiments involving cue competition effects (Beckers et al., 2009; Sobel & Kirkham, 2006, Experiment 1; Sobel et al., 2004), children were asked to generate types of actions they had never seen, involving the use of just one object that they had previously seen presented only as part of a pair. Indeed, Sobel and colleagues have argued that children's ability to generate a novel intervention suggests that in observing demonstrations children have learned something genuinely causal that goes beyond mere associative learning. They suggest that it demonstrates that children can use their inferences "to produce new interventions to elicit causal outcomes, rather than simply associating actions with effects" (Sobel et al., 2004, p. 328). In the context of the blicket detector task, the linguistic analogue of such a task would be to ask children to make judgments about situations they have not seen, in which one object, which they have so far seen only as part of a pair, is placed on the blicket detector on its own. That such an ability can be seen as a verbal reasoning analogue of generating a novel intervention is implied by recent ways of describing counterfactual judgments as involving the imagining of the effects of hypothetical interventions on a model of the causal structure of a scenario (Hagmayer et al., 2007; Sloman & Lagnado, 2005).

Experiment 1

The aim in this preliminary experiment was to replicate the findings of Sobel et al. (2004, Experiment 2), who found a difference between backward generative and blocking trials. In their experiment, Sobel et al. tested only 4-year-olds; in our study we also included an older group in order to examine whether the difference between trial types increased with age.

Method

Participants. Forty-five children took part in the study, nineteen 4-year-olds ($M = 54$ months, range = 49–59 months) and twenty-six 5- to 6-year-olds ($M = 74$ months, range = 69–80 months). There were 19 boys and 26 girls. In this experiment and subsequent ones, children were recruited from schools and nurseries and were tested in a quiet area in their schools. We recruited children in this and subsequent experiments by writing to the parents of all children in each school or nursery class that participated in the study. Children were included only if their parents agreed in writing to allow them to take part. In this and subsequent experiments, the children all were of Caucasian origin and came from English-speaking homes, and they were primarily of lower to middle social class. No specific information on parental income or educational level was available.

Materials. The "blicket detector" was constructed to be similar in appearance and operation to that used in previous studies by Gopnik and colleagues. It was made of gray plastic with a transparent red plastic top and was $20 \times 12 \times 7.5$ cm in dimension. A single wire emerged from the side of the box that was connected to a foot pedal hidden from the child's view. The experimenter could use the foot pedal to switch the box on and off without the child's knowledge. The box was battery powered. When the detector was switched on and an object was placed on top of it, a set of LEDs located under the red transparent top was immediately activated and a tune played. This was achieved through the use of weight-sensitive pressure pads. The lights turned off and the tune stopped as soon as the object was removed. When the detector was turned off, nothing happened when objects were placed on top of it. A pool of 30 wooden objects served as potential "blickets"; these were of varying colors, patterns, and shapes. This pool was kept in a box below the table on which the experiment was conducted.

Procedure. The pretraining procedure closely followed that of Sobel et al. (2004). Children were introduced to the blicket detector and told that it was a special machine that lit up and played a tune when blickets were placed on it. The experimenter then selected two objects from the pool and placed each one separately on the detector. The first object activated the detector; as children observed this the experimenter said, "This one makes the machine go. It is a blicket!" This object was then removed, and the second object was placed on the detector. It did not activate the detector, and the experimenter said, "This one doesn't make the machine go. It's not a blicket!" These demonstrations were then repeated. After this, children were told that they were going to help the experimenter play a game, which was to find out which objects were blickets. The experimenter took a further two objects from the pool for use in the training trials and placed them in front of the child. Each of the objects was then placed individually on the detector; one object set off the detector and the other object failed to do so. Children were asked whether or not each object was a blicket, and the two trials were then repeated. No children made errors on any of these training trials. Following the training phase, seven experimental trials were administered: two backward blocking trials, two backward generative trials, two association trials, and one basic control trial (see Table 2; the table also shows the expected patterns of responses). The experimenter randomly picked two objects from the pool before each trial and did not replace them

Table 2
Demonstrations Used in Each Trial Type

Trial type	Demonstration 1	Demonstration 2	Demonstration 3	Expected response to B
Association	A+	B+	B+	Yes
Forward				
Forward blocking	A+	AB+	AB+	Fewer yes responses than in forward blocking control, if blocking
Forward blocking control	A+	BC+	BC+	Yes or no
Forward generative	A−	AB+	AB+	Yes
Forward generative control	A−	BC+	BC+	Yes or no
Backward				
Backward blocking	AB+	AB+	A+	Fewer yes responses than in backward blocking control, if blocking
Backward blocking control	BC+	BC+	A+	Yes or no
Backward generative	AB+	AB+	A−	Yes
Backward generative control	BC+	BC+	A−	Yes or no

Note. AB denotes that A and B were placed simultaneously on the machine; BC denotes that B and C were placed simultaneously on the machine. Activation of the machine is indicated by a plus sign; nonactivation is indicated by a minus sign.

after the trial had been completed. The order in which children received the seven trials was randomized individually for each child. Children did not receive feedback during the test trials on whether or not their answers were correct.

Children received two backward blocking trials; in these trials children initially saw two demonstrations in which the paired objects (A and B) were placed on the detector, setting it off, followed by a demonstration in which one of the objects (A) set off the detector on its own. The left–right position of the A and B objects on the detector was counterbalanced; this was also the case for the backward generative trials. After the three demonstrations, the A and B objects were placed in front of the child, and, following Sobel et al. (2004), children were first asked whether or not the A object was a blicket and then whether or not the B object was a blicket. Children were encouraged to guess if they seemed unsure of an answer. Children were also given two backward generative trials; in these trials children initially saw two demonstrations in which A and B objects set off the detector when they were placed on it together, followed by a demonstration in which A was placed on the detector on its own but did not set it off.

There were also two trials, here referred to as association trials, that Sobel et al. (2004, Experiment 2) used in order to demonstrate that children are prepared to give positive responses to two cues in one trial. In these trials, there was one demonstration in which one object (A) was placed on the detector on its own and two demonstrations in which another object (B) was placed on the detector on its own. The detector went off on all three demonstrations. Children were asked if each object was a blicket. Children also received a single basic control trial in which three objects (A, B, and C) were used, with each placed on the detector individually. Two of these objects set the detector off and one did not; the object that set off the detector was varied across participants. Following Sobel et al., this basic control trial was included to ensure that children had understood the task and were paying attention to it. In this experiment, no children failed the basic control question. On completion of these trials, children were thanked for their participation and given a set of stickers of their choice.

Results and Discussion

Table 3 shows the average number of yes responses for each trial type (scores range from 0 to 2). In the case of all A objects, children could give a correct response simply by remembering whether or not the individually presented item had set off the detector. In both age–groups, participants invariably judged appropriately that the individually presented object was a blicket in the blocking trials and that it was not a blicket in the backward generative trials. They also invariably judged that both objects were blickets in the association trials.

For B objects, in backward generative trials the normatively correct response was to judge that B was a blicket. Indeed, all children consistently did so. In backward blocking trials, children had not been given enough information to judge whether or not B was a blicket, as although they had been shown that A was a blicket, it remained possible that B could be a blicket as well. Paired *t* tests showed that significantly fewer positive responses were given to B items in the backward blocking trials than in backward generative trials by the 5- to 6-year-olds, $t(25) = -4.67$, $p < .001$ ($M = -0.85$, 99% CI [−1.35, −0.34]), and the 4-year-olds, $t(18) = -4.40$, $p < .001$ ($M = -0.84$, 99% CI [−1.39, −0.29]), and that there was no significant difference between the age-groups in responses to blocking trials, $t(43) = -.02$, $p = .99$

Table 3
Mean Number of Yes Responses for Each Trial Type in Experiment 1

Trial type	4-year-olds A	4-year-olds B	5- to 6-year-olds A	5- to 6-year-olds B
Backward blocking	2.00 (0.00)	1.15 (0.83)	2.00 (0.00)	1.16 (0.92)
Backward generative	0.00 (0.00)	2.00 (0.00)	0.00 (0.00)	2.00 (0.00)
Association	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)	2.00 (0.00)

Note. Standard deviations are given in parentheses.

($M = -0.004$; 95% *CI* [−0.54, 0.54]). These findings are essentially a replication of those of Sobel et al. (2004, Experiment 2).

Experiment 2

In Experiment 1, participants' responses to the paired B objects were compared across blocking trials and generative trials. However, as we omitted the appropriate controls in order to replicate Sobel et al. (2004, Experiment 2) it is not clear whether such findings reflect backward blocking, generative learning, or both (Beckers, Van den Broeck, et al., 2005; Beckers et al., 2009). In our second experiment we used the appropriate control trials. In these control trials, participants initially saw two objects placed together on the detector, as in backward generative and backward blocking trials. However, for the subsequent demonstration of an individually presented object, a new object was introduced that had not been part of the pair (see Table 2). In the controls for backward blocking trials the new object set off the detector; in the controls for backward generative trials it did not. Because the object was new, these demonstrations with an individually presented object should not have led to a revaluation of the objects previously presented in a pair. Differences between responses to the B object in these control trials in comparison to those in the backward blocking or generative test trials can be confidently interpreted as an indication that in the test trials the causal status of B has been revaluated as a result of the subsequent presentation of A on its own. In this experiment we also introduced a forward condition, in which demonstrations involving the individually presented objects were given before the paired presentations. Our aim was to compare levels of blocking in this forward condition to those observed in the backward condition. As mentioned previously, existing theoretical accounts of causal learning differ fundamentally in terms of whether or not they predict an effect of order of presentations (Cheng, 1997; Dickinson, 2001; Kruschke, 2006; Shanks, 2006). For example, although simple associative theories predict forward blocking, they do not predict its backward equivalent. However, many nonassociative theories do not readily predict an effect of order of presentations, as learning is based on pieces of information that are identical regardless of trial order (e.g., Cheng, 1997; Gopnik et al., 2004; for a discussion of this point, see Shanks, 2006). Thus, of interest is whether forward blocking and backward blocking are observed at similar levels and whether they show similar developmental patterns.

Method

Participants. Ninety-three children took part in the study: forty-three 4-year-olds and fifty 5- to 6-year-olds. Five of the 4-year-olds and four of the 5- to 6-year-olds failed the basic control question, and they were not included in further analyses. Of the remaining 4-year-olds, 19 took part in the forward condition (M = 53 months, range = 48–59 months) and 19 took part in the backward condition (M = 53 months, range = 48–58 months); of the 5- to 6-year-olds, 23 took part in the forward condition (M = 76 months, range = 69–80 months) and 23 took part in the backward condition (M = 74 months, range = 69–79 months). There were 42 boys and 42 girls.

Materials. These were identical to those used in Experiment 1.

Procedure. Children were randomly assigned to either the forward or the backward condition. The order in which children received the trials was randomized individually for each child. In the backward condition children received two each of four types of trials (see Table 2) plus one additional basic control trial. The two backward blocking trials and backward generative trials were identical to those in Experiment 1. The control trials were identical to experimental trials except that the object placed on the blicket detector alone was not one of those previously placed on the detector in a pair. The trials in the forward condition were identical to those in the backward condition except that all the individual presentations of the A object occurred before the paired demonstrations had been shown.

Results

Table 4 shows the average number of "yes" responses (with scores ranging from 0 to 2) given to A and B items in each trial type as a function of age-group and condition.

Responses to A. An initial four-way analysis of variance (ANOVA) was conducted on the number of yes responses to A items with factors of age-group (4-year-olds vs. 5- to 6-year-olds), condition (forward vs. backward), trial type (blocking vs. generative), and trial status (control vs. experimental). As expected, the

Table 4
Mean Number of Yes Responses Out of Two for Each Trial Type in Experiment 2

Condition and trial type	4-year-olds		5- to 6-year-olds	
	A	B	A	B
Forward				
Forward blocking	2.00 (0.00)	1.89 (0.32)	1.91 (0.29)	1.13 (0.97)
Forward blocking control	1.79 (0.54)	1.94 (0.23)	1.82 (0.49)	1.57 (0.73)
Forward generative	0.74 (0.73)	2.00 (0.00)	0.26 (0.54)	1.87 (0.34)
Forward generative control	0.26 (0.45)	1.94 (0.23)	0.13 (0.34)	1.57 (0.66)
Backward				
Backward blocking	2.00 (0.00)	1.84 (0.37)	2.00 (0.00)	1.00 (0.90)
Backward blocking control	2.00 (0.00)	1.89 (0.57)	2.00 (0.00)	1.52 (0.67)
Backward generative	0.05 (0.23)	1.95 (0.23)	0.00 (0.00)	2.00 (0.00)
Backward generative control	0.16 (0.50)	1.74 (0.65)	0.04 (0.21)	1.61 (0.66)

Note. Standard deviations are given in parentheses.

main effect of trial type was highly significant, $F(1, 80) = 1{,}654.35$, $p < .001$, $\eta_p^2 = .95$, with many more yes responses given to A objects in blocking trials than in generative trials. There were a number of significant two-way interactions. There was a two-way interaction between trial status and condition, $F(1, 80) = 10.26$, $p < .01$, $\eta_p^2 = .11$; further analysis showed that the effect of trial status was significant in the forward, $F(1, 41) = 8.90$, $p < .01$, $\eta_p^2 = .16$, but not the backward conditions, $F(1, 41) = 1.30$, $p = .26$, $\eta_p^2 = .03$, with significantly more yes responses in the experimental than in the control trials in the forward condition. There was also a significant interaction between trial type and age-group, $F(1, 80) = 4.52$, $p = .037$, $\eta_p^2 = .05$; further analyses showed that the effect of age-group was significant for generative trials, $t(82) = 2.71$, $p < .01$ ($M = 0.19$, 99% CI [0.01, 0.38]), but not for blocking trials, $t(82) = 0.27$, $p = .79$ ($M = 0.01$, 99% CI [−0.11, 0.14]), with the younger age-group producing more erroneous yes responses in generative trials than the older age-group. Last, there was also a significant interaction between trial type and condition, $F(1, 80) = 22.21$, $p < .001$, $\eta_p^2 = .22$. Further analyses showed that performance was generally poorer in the forward condition, with participants giving more erroneous yes responses in generative trials in this condition than in the backward condition, $F(1, 82) = 16.27$, $p < .001$, $\eta_p^2 = .17$, and fewer correct yes responses to blocking trials in this condition than in the backward condition, $F(1, 82) = 7.17$, $p < .05$, $\eta_p^2 = .08$. There were no other significant interactions.

In summary, although for both age-groups the vast majority of responses to A objects were correct, performance was significantly worse in the forward than the backward condition. In the former condition, the individual presentations occurred before the paired presentations rather than after, and it seems likely that the effect of condition on answers to A is a result of participants simply being more likely to forget what had happened when A had been demonstrated on its own. The younger group also made significantly more erroneous yes responses in generative trials than did the older group, although there were no age effects on the number of correct yes responses in blocking trials.

Responses to B. In this experiment, blocking is indicated only if participants give fewer yes responses to B in blocking trials than in blocking control trials. Correspondingly, generative learning is indicated only if participants give more yes responses to B in these trials than in the respective control trials. It can be seen from Table 4 that the proportions of yes responses to B objects in experimental trials appear to be different from those in control trials only in the older age-group.

A four-way ANOVA was conducted on the number of yes responses to B objects with factors of age-group (4-year-olds vs. 5- to 6-year-olds), condition (forward vs. backward), trial type (blocking vs. generative), and trial status (control vs. experimental). The two-way interaction between trial type and age-group was significant, $F(1, 80) = 15.05$, $p < .001$, $\eta_p^2 = .16$, as was the three-way interaction among trial type, trial status, and age-group, $F(1, 80) = 10.43$, $p < .01$, $\eta_p^2 = .12$, whereas the main effect of condition was not significant ($F < 1$), and there were no significant two- or three-way interactions between condition and any other factors (Fs < 1.3). Thus, further analyses were conducted separately for each age-group collapsed across conditions. In the younger group, there was a significant interaction between trial type and trial status, $F(1, 37) = 6.19$, $p < .05$, $\eta_p^2 = .14$. However, paired t tests showed that there was no difference between experimental and control trials either for blocking trials, $t(37) = -0.57$, $p = .57$ ($M = -0.03$, 99% CI = [−0.15, 0.10]), or for generative trials, $t(37) = 1.71$, $p = .10$ ($M = 0.13$, 99% CI = [−0.08, 0.34]). For the older group, there was also a significant interaction between trial type and trial status, $F(1, 45) = 24.10$, $p < .001$, $\eta_p^2 = .35$. Paired t tests showed that there was a significant difference between experimental and control trials for both blocking trials, $t(45) = -3.76$, $p < .001$ ($M = -0.47$, 99% *CI* [−.82, −.14]), and generative trials, $t(45) = 3.5$, $p < .01$ ($M = 0.35$, 99% CI [0.08, 0.62]).

Discussion

The analyses indicate that cue competition effects occurred only in the older group, as it was only in this group that experimental trials differed from control trials. The effect of condition was not significant, indicating that the order in which participants received the demonstrations had no effect on their performance. However, clear age effects were found. The fact that 4-year-olds did not demonstrate either blocking or generative learning in this task may lead us to rethink Sobel et al.'s (2004) interpretation of the type of trial difference found in Experiment 1. We have argued that a comparison of blocking trials with the type of control trial used in Experiment 2 is more appropriate than a comparison of blocking trials with generative trials, as our control trials align more closely with those used in the literature on adult causal learning. Nevertheless, it is of interest that for the 4-year-old group, the mean number of positive responses for B objects in backward blocking trials found in Experiment 2 is notably higher than that found in Experiment 1. Indeed, we note that the responses of most of the 4-year-olds (30 out of 38) in Experiment 2 are consistent with their adopting the heuristic that all objects are blickets except those that they have seen not setting off the detector. In fact, in this respect the responses of our 4-year-olds closely resembled those of the 3-year-olds in Beckers et al.'s (2009) study, who also tended to give positive responses unless they had directly observed that an object did not set off the detector. The 5-year-old group, by contrast, produced somewhat fewer positive responses to B objects in Experiment 2 than in Experiment 1.

Why might 4-year-olds have been more likely to judge that B was not a blicket in backward blocking trials in Experiment 1? One possible interpretation of the findings of Experiment 1 is that children believed they were being asked to decide which of the two objects was a blicket, and if they had judged that A was a blicket, they would by default judge that B was not a blicket in blocking trials (and conversely in generative trials). Sobel et al. (2004) argued that the inclusion of association trials controls for such a possibility. However, it may be the case that children treat the questions as forced choice only when they have seen demonstrations of paired objects, whereas items are always demonstrated individually in association trials. Our suggestion is that negative responses in blocking trials in Experiment 1 may have reflected a simple tendency to treat the two questions about A and B as forced choice (i.e., that children took their task to be to decide which one of two items was the blicket). Thus, the issue is why children did not adopt this approach in Experiment 2. One likely explanation is that the presence of the control trials discouraged such a tendency. The inclusion of these trials may have meant that children were less likely to believe that their task was to make a forced choice between A and B object, because these objects never appeared together in these trials.

In conclusion, on the basis of the findings of Experiment 2, we argue that the findings of Experiment 1 and those of Sobel et al. (2004) give us no reason to believe that 4-year-olds show genuine blocking on this type of task. In this respect, our results are compatible with those of Beckers et al. (2009). This finding is important because it suggests that whatever account we give of the processes underlying blocking on this task needs to be at least amenable to the possibility that these processes change developmentally. With respect to generative trials, our results are also identical to the findings of Beckers et al. (2009), who found that 3-year-olds were equally likely to classify as a blicket the relevant object from a backward generative pair and one from control pair. However, we note that Beckers et al. did find that children were more likely to correctly choose the relevant object from a backward generative pair, rather than from the corresponding control pair, when asked to choose an object to set off the detector. This raises the possibility that response mode affects performance in this type of task, an issue we return to the General Discussion.

Experiment 3

In Experiment 3, we asked children to consider scenarios that differed from the ones they had observed, in order to examine whether they could bring their causal knowledge to bear in answering counterfactual questions. We used the same trial types to assess cue competition effects as those used in Experiment 2 (e.g., AB+, A+ backward blocking trials), but rather than asking children whether or not the B object was a blicket, we asked them whether the detector would have gone off in a counterfactual scenario in which A had not been placed on the detector along with B. Other researchers have used counterfactual questions as a measure of causal inference with some success (e.g., Buehner, Cheng, & Clifford, 2003; Collins & Shanks, 2006), although the tasks and the counterfactual questions used with adults are considerably more complex than the one we used in this study with children. It is safe to assume that adults can grasp the counterfactual scenarios referred to in such questioning, but we do not know whether even 5- to 6-year-old children are capable of the type of counterfactual reasoning that has been described. Although children as young as 3 can make some types of counterfactual judgments (e.g., German & Nichols, 2003; Harris et al., 1996), age differences are evident in other types of counterfactual tasks (Beck et al., 2006; German & Nichols, 2003; Perner et al., 2004; Riggs, Peterson, Robinson, & Mitchell, 1998). Moreover, most previous studies of children's counterfactual reasoning have used simple scenarios involving causal relations with which young children might be familiar (e.g., a doll walking in muddy shoes that make the floor dirty) rather than situations in which children have to make inferences about the entirely novel causal powers of unfamiliar objects. This made it difficult to predict on the basis of existing research findings whether or not young children would be able to answer our counterfactual questions.

Method

Participants. Twenty-one 4-year-olds (M = 52 months, range = 49–56 months) and thirty 5- to 6-year-olds (M = 69 months, range = 63–75 months) took part in this experiment. There were 22 boys and 29 girls.

Materials. These were identical to those used in the previous experiments.

Procedure. The introduction to the machine and the training trials were identical to those in Experiment 2. At test, children received two trials of each of the four backward trials used in Experiment 2 (see Table 2). Following the last demonstration in each trial, children were asked two questions. They were first asked whether or not A was a blicket, while the experimenter held A in one hand. Following the answer to this question, in backward blocking and generative trials, the experimenter picked up B with her other hand, so that she was holding both object A and object B in front of the child, with one object in each hand. Children were then asked a counterfactual question: The experimenter said, "A moment ago I put both these blocks on the machine together, and it went off. Do you think it would have gone off if I hadn't put this [gave color of A] one on?" The color of the A object was referred to in the question to make completely sure children knew to which object the experimenter was referring, because the experimenter was holding both objects. As she was asking the question, the experimenter also moved the hand holding the A object toward herself and away from the child to make it clear which object she meant, leaving the hand holding the B object in front of the child. In control trials, children were also initially asked whether A was a blicket. However, A did not feature in the counterfactual question. In these control trials, the experimenter put A aside (following the initial question about whether or not it was a blicket) and then picked up objects B and C. As in the other trial types, the experimenter then said, "A moment ago I put both these blocks on the machine together, and it went off. Do you think it would have gone off if I hadn't put this [gave color of C] one on?"

Results and Discussion

Table 5 shows the proportion of positive responses given to each trial type. It can be seen from the table that children gave the highest number of positive responses to the counterfactual question in backward generative experimental trials. In backward blocking trials, the response consistent with judging that B was not a blicket would be to give a negative answer to the counterfactual. Five- to six-year-olds gave the smallest number of positive responses in these trials. The corresponding control trials give a baseline measure of the tendency to give positive responses to paired objects that set off the detector; it can be seen that children produced equal numbers of positive responses to the two control trial types.

Table 5
Mean Number of Positive Responses Out of Two Given to Counterfactual Questions in Experiment 3

Trial type	4-year-olds	5- to 6-year-olds
Backward blocking	1.48 (0.75)	0.70 (0.84)
Backward blocking control	1.48 (0.81)	0.97 (0.81)
Backward generative	1.57 (0.68)	1.50 (0.73)
Backward generative control	1.33 (0.80)	1.00 (0.87)

Note. Standard deviations are given in parentheses.

A three-way ANOVA on the number of positive responses given to the counterfactual question with a between-subjects factor of age and within-subjects factors of trial type (blocking vs. generative) and trial status (experimental vs. control) was conducted. The main effect of group was significant, $F(1, 49) = 5.23$, $p < .05$, $\eta_p^2 = .096$, with the 4-year-olds giving more positive responses than did the older children. The interaction between trial type and trial status was significant, $F(1, 49) = 8.35$, $p < .01$, $\eta_p^2 = .146$, as was the interaction between trial type and age, $F(1, 49) = 9.19$, $p < .01$, $\eta_p^2 = .158$. Separate two-way analyses of variance were conducted for each age-group; the interaction between trial type and trial status was significant only in the 5- to 6-year-old group, $F(1, 29) = 8.23$, $p < .01$, $\eta_p^2 = .223$. Paired *t* tests showed that none of the differences between any pair of trial types were significant in the 4-year-old group (*p*s $> .05$). In the 5- to 6-year-old group, paired *t* tests found significant differences between backward blocking and backward generative trials, $t(29) = -4.56$, $p < .001$ ($M = -0.80$, 99% CI [-1.28, -0.32]), and between backward generative trials and their control trials, $t(29) = 3.04$, $p < .001$ ($M = 0.50$, 99% CI [0.05, 0.95]). However, although the difference between backward blocking and backward blocking control trials was in the predicted direction it failed to reach significance, $t(29) = -1.68$, $p = .10$ ($M = -0.27$, 99% CI [-0.70, 0.17]).

These results are largely consistent with those in Experiment 2, although overall levels of positive responses were markedly lower for both groups in this task. The pattern of counterfactual judgments of the 4-year-olds closely followed the pattern of causal judgments found for this age-group in Experiment 2, with similar responses given for the different trial types. Thus, for this age-group, again, no evidence for cue competition effects was found. The pattern of counterfactual judgments given by the older group was similar to that found for causal judgments in Experiment 2, in that these children gave different types of judgments for the different trial types: Children were more likely to judge that the detector would not have gone off without A being present in blocking trials than in generative trials. However, the failure of the difference between backward blocking experimental trials and control trials to reach significance means that the results do not exactly match those of Experiment 2.

Experiment 4

In Experiment 3, children were asked the counterfactual question "Do you think the machine would have gone off if I hadn't put this one [A] on?" Arguably, it is only implicit within the question that children should be considering a counterfactual alternative in which B was placed on the detector, despite A's absence. In Experiment 4, we asked children the closely related counterfactual question "Do you think the machine would have gone off if I had only placed B on it?" It is possible that making the counterfactual alternative more explicit might assist children; furthermore this question focuses children's attention on the causal powers of the item of interest, namely B. We predicted that the same general pattern of age-related differences should be preserved, although differences between the different trial types might be magnified if the form of questioning assists children to imagine and focus on the relevant counterfactual scenario.

Method

Participants. Twenty-one 4-year-olds ($M = 52$ months, range = 49–58 months) and twenty-six 5- to 6-year-olds ($M = 74$ months, range = 64–83 months) took part in this experiment. There were 23 boys and 24 girls.

Materials. These were identical to those used in the previous experiments.

Procedure. This was identical to that in Experiment 3 except in the nature of the test questions. Children received the same eight trial types as in Experiment 3; following the last demonstration children were again asked whether or not the A item was a blicket. As in Experiment 3, they were then asked a counterfactual question, but in this experiment they were asked to judge what would have happened if B had been placed on the detector on its own. The experimenter said, "A moment ago I put both of these blocks on the machine together, and it went off. Do you think it would have gone off if I had only put this [gave color of B] one on?"

Results and Discussion

Table 6 shows the average number of positive responses given to each trial type as a function of age-group. A three-way ANOVA on the number of positive responses given to the counterfactual question with a between-subjects factor of age and within-subjects factors of trial type (blocking vs. generative) and trial status (experimental vs. control) was conducted. The main effect of age was significant, $F(1, 45) = 4.86$, $p < .05$, $\eta_p^2 = .098$, with the 4-year-olds giving more positive responses than did the older children. The two-way interaction between trial type and trial status was significant, $F(1, 45) = 9.05$, $p < .01$, $\eta_p^2 = .167$, as was the three-way interaction among age, trial type, and trial status, $F(1, 45) = 5.76$, $p < .05$, $\eta_p^2 = .113$. Given this, separate two-way ANOVAs were conducted for each age-group. The interaction between trial type and trial status was significant in the older group, $F(1, 25) = 13.94$, $p < .01$, $\eta_p^2 = .358$, but not in the younger group, $F(1, 20) = 0.21$, $p > .05$. Paired *t* tests showed that in the older group there were significant differences between backward blocking trials and backward blocking control trials, $t(25) = -2.52$, $p < .02$ ($M = -0.42$, 95% CI [-0.77, -0.08]); backward generative trials and backward generative control trials, $t(25) = 2.67$, $p < .02$ ($M = 0.42$, 95% CI [0.10, 0.75]); and backward blocking and backward generative trials, $t(25) = -2.59$, $p < .02$ ($M = -0.58$, 95% CI [-1.01, -0.12]). However, none of these differences between trial types were significant in the 4-year-old group (*p*s $> .10$).

Table 6
Mean Number of Positive Responses Out of Two Given to Counterfactual Questions in Experiment 4

Trial type	4-year-olds	5- to 6-year-olds
Backward blocking	1.38 (0.92)	0.96 (0.92)
Backward blocking control	1.57 (0.75)	1.38 (0.64)
Backward generative	1.67 (0.58)	1.54 (0.65)
Backward generative control	1.76 (0.54)	1.12 (0.71)

Note. Standard deviations are given in parentheses.

These results are consistent with the findings from Experiment 2 and 3, in that no evidence of cue competition effects was found in the younger age-group. In this experiment, the pattern of results for the older group closely matched those for the equivalent age-group in Experiment 2, in which children were asked causal rather than counterfactual questions. In contrast to the findings of Experiment 3, the difference between blocking trials and control trials was significant for the older group. This suggests that the older children are capable of making the relevant counterfactual judgments, particularly when the nature of the counterfactual scenario they are asked to consider is made clear.

General Discussion

Our experiments examined a number of different cue competition effects in young children in the context of the blicket detector task, and their results can be summarized as follows. First, we found clear evidence of age differences: Cue competition effects were reliably observed in our 5- to 6-year-old group but not in our 4-year-old group (Experiments 2–4). Second, we found no effect of trial order, in that cue competition effects that involved backward designs were no more difficult to obtain than those that did not (Experiment 2). Third, we found that blocking effects were as easily obtained as generative effects, and this was the case even for the effects in backward procedures (Experiments 2 and 4). Last, backward blocking and generative effects were also obtained in our older group when children were asked a more complex counterfactual question about what had happened in the absence of a cue, rather than simply asked to make a categorization judgment (Experiments 3 and 4).

Age Effects

We have interpreted the results of Experiment 2, further supported by Experiments 3 and 4, as demonstrating a developmental change between four and five years in causal judgments. We note that other researchers have found other sorts of developmental changes in causal judgments using this type of procedure (Gopnik et al., 2001; Sobel & Kirkham, 2006; Sobel et al., 2004). What we take to be of interest here is not the specific age at which cue competition effects were found to be emerging (which, after all, may differ in other populations) but the fact that robust age-related improvements were obtained, suggesting that we need an account of such effects that allows for, or even predicts, such improvements. If it is the case, as De Houwer (2009; De Houwer et al., 2005) has argued, that effortful controlled reasoning processes that load on working memory underpin cue competition effects in humans, such age differences are entirely unsurprising.

Having said that, we note that we are not arguing here that cue competition effects could never be observed in 4-year-olds. Beckers, Van den Broeck, et al.'s (2005) study provides at least some evidence to the contrary; furthermore, the fact that such phenomena are widely observed in animals provides good reason to believe that under some circumstances it may be possible to demonstrate cue competition effects even in very young children (see also Sobel & Kirkham, 2006). Rather, as many researchers have argued (see, e.g., Lovibond et al., 2003; Pineňo & Miller, 2007; Vadillo & Matute, 2007), it seems likely that there may not be a single set of processes underpinning phenomena such as blocking across all circumstances. Which processes are involved may very much depend upon the nature of the task itself. Our findings raise the challenge of explaining why such a developmental pattern may emerge in a procedure that has been used very extensively in recent studies of children's causal learning. Elsewhere, it has been suggested that it is particularly likely that explicit inferential reasoning processes rather than more basic associative mechanisms may be important in tasks in which there are relatively few cues and low numbers of presentations (De Houwer et al., 2005; see also Le Pelley & McLaren, 2003). Given that the current task employs the minimal number of cues/presentations possible to demonstrate cue competition phenomena, it is particularly plausible that the former sort of processes play a role. Their involvement may be used to explain developmental patterns on this task.

Cue Competition Effects

Whatever account is given of such phenomena on this task, it must be one that does not predict that the order in which participants receive information about stimuli affects the conclusions participants reach about their causal status. This finding is noteworthy because with adults it has sometimes proved more difficult to demonstrate backward cue competition effects, such as backward blocking, than their forward equivalents. However, although backward learning has sometimes been viewed as primarily a feature of potentially more sophisticated human causal learning, under at least some conditions related effects have been observed in animal learning (e.g., Balleine, Espinet, & Gonzalez, 2005; Liljeholm & Balleine, 2006; Miller & Matute, 1996). The finding of equivalent levels of cue competition effects across our forward and backward conditions (Experiment 2) might be taken to suggest that simple associative accounts of blocking are inadequate for this task (see also Sobel et al., 2004), although we emphasize strongly that such associative accounts were not designed to model findings from this sort of paradigm, in which a very small number of cue–outcome pairings are demonstrated.

In general, we found that backward blocking was as readily observable as backward generative learning. Again, this finding is of note, because backward blocking effects have sometimes proved more elusive to demonstrate than backward generative learning (see Dickinson, 2001) and the two effects are differentially affected by some task variables (Beckers, De Houwer, et al., 2005; Lovibond et al., 2003; Vandorpe & De Houwer, 2005). Furthermore, Beckers et al. (2009) claimed to have found evidence for the latter but not the former in their 3-year-old group. Indeed, Beckers et al. argued that, for this particular task, higher order accounts predict generative learning but not blocking. Beckers et al. provided a more detailed discussion of why they think this; effectively, they argued that participants have not been provided with sufficient information to be confident about B's status. Although we agree that children could not definitively conclude B was not a blicket, we nevertheless believe that a sensitivity to the differential probability that a cue is a blicket across blocking and control trials could result in a difference in responses to cues across these trial types. Recall that the measure of blocking is given by the difference between causal judgments given to, for example, B after A+ and AB+ demonstrations in comparison to judgments given to C after CD+ demonstrations. Normatively, the likelihood that B is causal under such circumstances is identical to the baseline probability that any given cue is causal (see also Sobel & Kirkham, 2007; Sobel et al., 2004; Tenenbaum & Griffiths, 2003); in effect, participants have been provided with no useful information about B's status. However, the likelihood that C is causal will always be higher than the baseline unless the baseline probability is 1. Of course, the probability that C is causal is also related to the baseline and can be derived from Bayes' law to be 1/(2 − baseline); thus, for example, if the baseline probability is .50, the probability that C is causal is .66. If participants are sensitive to the fact that these probabilities are different, we would expect to see a difference in judgments to B and C. Thus, we take our findings to suggest that at least by age 5–6 years children are sensitive to the differing probabilities between these two trial types. Note that this sensitivity need not take the form of children actually calculating, even implicitly, the specific probabilities that obtain in a given task (after all, most adults would find it hard to derive the necessary probability calculations). Rather, the minimum that is required is that their judgments track the fact that in blocking trials they have effectively obtained no information regarding B's status, whereas in control trials they have received some positive information that a cue is causal.

Counterfactual and Causal Judgments

Experiments 3 and 4 introduced a novel way of assessing children's causal inferences on this task. Rather than being asked whether or not an object was a blicket, children were asked counterfactual questions about whether or not the detector would have gone off in the absence of one of the objects from a pair. Answering such questions required children to think about a scenario they had never observed: a scenario in which an object, which they had seen on the detector only along with another object, was placed on the detector individually. Under this form of questioning, our oldest group again showed evidence of cue competition effects, particularly when the counterfactual scenario to be considered was made very clear (Experiment 4). To our knowledge, these studies represent the first attempts to examine young children's counterfactual judgments in a task in which they have to make inferences about whether or not unfamiliar objects possess a completely novel causal property. It has previously been argued that some methods of assessing young children's counterfactual reasoning may have failed to properly tap counterfactual cognition, because children could have answered the questions correctly on the basis of preexisting knowledge without engaging in genuinely counterfactual thought about the experimental scenarios (see Perner et al., 2004). Such a criticism cannot apply with regard to the questioning in our study, because children had to answer questions about scenarios they had never observed and could not have possessed the requisite preexisting knowledge. However, there may be other ways in which young children could answer counterfactual questions correctly without engaging in genuine counterfactual thought (Beck et al., 2006). For example, children might translate the questions as being about what would happen if a certain event were to occur in the future (i.e., future hypothetical thinking: What will happen if B is placed on the detector on its own?) rather than what would have happened if things had been different (i.e., counterfactual thinking: What would have happened if B had been placed on the detector on its own?). There is some evidence that reasoning hypothetically about the future may emerge developmentally earlier than counterfactual reasoning (see Beck et al., 2006; Riggs et al., 1998). We note that our older participants could have shown blocking or generative learning on this task only if they had taken into consideration the fact that B had been demonstrated along with A, so they must have at least considered past events that they had just observed. However, our experimental procedure does not allow us to be completely sure that children did not then switch to thinking about a future hypothetical situation rather than a specifically counterfactual one.

Nevertheless, introducing this method of assessing children's causal inferences means that three methods have now been employed to assess children's abilities on the much-used blicket detector task: asking children directly whether or not objects are blickets (i.e., categorization judgments); asking them to make novel interventions (e.g., "make it go" or "make it stop" by placing objects on or taking them off the blicket detector); and, now, asking them to make counterfactual judgments. Gopnik and colleagues have argued that the ability to make the second type of judgment indicates that children who are categorizing objects as blickets do so on the basis of genuinely causal rather than merely associative learning (e.g., Gopnik et al., 2001; Sobel et al., 2004). Furthermore, we argued in the Introduction that the third type of judgment, counterfactual judgment, can potentially be seen as the verbal reasoning analogue of asking children to make a novel intervention. In our studies, cue competition effects showed a similar developmental timetable when children were asked either categorization questions (Experiment 2) or counterfactual questions (Experiment 3); however, we did not present both types of questions to a given sample of children, and none of our experiments employed the action-based response measure. Thus, we cannot yet fully identify the interrelationships between these three types of judgments: It may be that these types of abilities dissociate developmentally, with phenomena such as cue competition effects appearing earlier when purely action-based response are used, or it may be that all of these response measures tap a common process and that children who are competent with one response measure will show competence using the other measures. We are currently exploring this issue in our empirical work.

Although the suggestion that making counterfactual judgments forms part of the process of arriving at causal judgments (e.g., Harris et al., 1996) has largely fallen out of favor within experimental psychology (e.g., German, 1999; Mandel, 2003), there is still considerable debate over the nature of the relation between these two types of judgment. In particular, it has recently been argued by those who describe causal learning in terms of the construction of models of the relations between variables, such as Gopnik et al. (2004) and Waldmann et al. (2006), that the sorts of representations that support causal judgments should also be expected to support counterfactual judgments (see Hagmayer et al., 2007; Sloman & Lagnado, 2005). There is clearly a need for researchers to examine the relation between children's causal and counterfactual judgments to explore whether or not this is the case. Our results make some contribution to this debate, insofar as they suggest that whatever account is provided of the processes and/or representations underpinning children's causal learning on the blicket detector task must be one that can explain children's ability to make not only causal judgments but also counterfactual ones.

References

Aitken, M. R. F., Larkin, M. J. W., & Dickinson, A. (2001). Reexamination of the role of within-compound associations in the retrospective revaluation of causal judgments. *Quarterly Journal of Experimental Psychology, 54*B, 27–51.

Aloise, P. A., & Miller, P. H. (1991). Discounting in preschoolers: Effect of type of reward agent. *Journal of Experimental Child Psychology, 52,* 70–86.

Balleine, B. W., Espinet, A., & Gonzalez, F. (2005). Perceptual learning enhances retrospective revaluation of conditioned flavor preferences in rats. *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 341–350.

Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77,* 413–426.

Beckers, T., De Houwer, J., Pineño, O., & Miller, R. R. (2005). Outcome additivity and outcome maximality influence cue competition in human causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 238–249.

Beckers, T., Van den Broeck, U., Renne, M., Vandorpe, S., De Houwer, J., & Eelen, P. (2005). Blocking is sensitive to causal structure in 4-year-old and 8-year-old children. *Experimental Psychology, 52,* 264–271.

Beckers, T., Vandorpe, S., Debeys, I., & De Houwer, J. (2009). Three-year-olds' retrospective revaluation in the blicket detector task: Backward blocking or recovery from overshadowing? *Experimental Psychology, 56,* 27–32.

Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1119–1140.

Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 837–854.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition, 18,* 537–545.

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

Collins, D. J., & Shanks, D. R. (2006). Conformity to the power PC theory of causal induction depends on the type of probe question. *Quarterly Journal of Experimental Psychology, 59,* 225–232.

Collins, J., Hall, N., & Paul, L. A. (Eds.). (2004). *Causation and counterfactuals.* Cambridge, MA: MIT Press.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior, 37,* 1–20.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental Psychology, 55*B, 289–310.

De Houwer, J., & Beckers, T. (2003). Secondary task difficulty modulates forward and backward blocking in human contingency learning. *Quarterly Journal of Experimental Psychology, 56*B, 345–357.

De Houwer, J., Beckers, T., & Vandorpe, S. (2005). Evidence for the role of higher-order processes in cue competition and other learning phenomena. *Learning & Behavior, 33,* 239–249.

Denniston, J. C., Savastano, H. I., & Miller, R. R. (2003). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65–117). Mahwah, NJ: Erlbaum.

Dickinson, A. (2001). Causal learning: An associative analysis. *Quarterly Journal of Experimental Psychology, 54*B, 3–25.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology, 49*B, 60–80.

German, T. (1999). Children's causal reasoning: Counterfactual thinking occurs for "negative" outcomes only. *Developmental Science, 2,* 442–447.

German, T., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science, 6,* 514–523.

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.

Gopnik, A., & Schulz, L. E. (2004). Mechanisms of theory formation in young children. *Trends in Cognitive Sciences, 8,* 371–377.

Gopnik, A., & Schulz, L. E. (2007). Introduction. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 1–15). Oxford, England: Oxford University Press.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37,* 620–629.

Hagmayer, Y., Sloman, S., Lagnado, D., & Waldmann, M. (2007). Causal reasoning through intervention. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 86–100). Oxford, England: Oxford University Press.

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition, 61,* 233–259.

Hart, H. L. A., & Honoré, T. (1959/1985). *Causation in the law* (2nd ed.). Oxford, England: Clarendon.

Kelley, H. H. (1972). *Causal schemata and the attribution process.* New York, NY: General Learning Press.

Kruglanski, A. W., Schwartz, S. M., Maides, S., & Hamel, I. Z. (1978). Covariation, discounting, and augmentation: Towards a clarification of attributional principles. *Journal of Personality, 76,* 176–189.

Kruschke, J. K. (2006). Local Bayesian learning with applications to retrospective revaluation and highlighting. *Psychological Review, 113,* 677–699.

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychological Bulletin & Review, 7,* 636–645.

Kushnir, T., & Gopnik, A. (2005). Young children infer causal strength from probabilities and interventions. *Psychological Science, 16,* 678–683.

Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43,* 186–196.

Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *Quarterly Journal of Experimental Psychology, 56*B, 68–79.

Le Pelley, M. E., Oakeshott, S. M., & McLaren, I. P. L. (2005). Blocking and unblocking in human causal learning. *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 56–70.

Lewis, D. (1973). Causation. *Journal of Philosophy, 70,* 556–567.

Liljeholm, M., & Balleine, B. W. (2006). Stimulus salience and retrospective revaluation. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 481–487.

Lovibond, P. F., Been, S., Mitchell, C. J., Bouton, M. E., & Frohardt, R. (2003). Forward and backward blocking of causal judgment is enhanced by additivity of effect magnitude. *Memory & Cognition, 31,* 133–142.

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

Mandel, D. R. (2003). Judgment dissociation theory: An analysis of differences in causal, counterfactual, and covariational reasoning. *Journal of Experimental Psychology: General, 132,* 419–434.

Miller, P. H., & Aloise, P. A. (1989). Young children's understanding of the psychological causes of behavior: A review. *Child Development, 60,* 257–285.

Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General, 125,* 370–386.

Mitchell, C. J., Lovibond, P. F., & Condoleon, M. (2005). Evidence for deductive reasoning in blocking of causal judgements. *Learning and Motivation, 36,* 77–87.

Perner, J., Sprung, A. R., & Steinkogler, B. (2004). Counterfactual conditionals and false belief: A developmental dissociation. *Cognitive Development, 19,* 179–201.

Pineňo, O., & Miller, R. R. (2007). Comparing associative, statistical, and inferential reasoning accounts of human contingency learning, *Quarterly Journal of Experimental Psychology, 60,* 310–329.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development, 13,* 73–90.

Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared cause tummy aches? Naive theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology, 43,* 1124–1139.

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology, 37*B, 1–21.

Shanks, D. R. (2006). Bayesian associative learning. *Trends in Cognitive Sciences, 10,* 477–478.

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 60,* 291–309.

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

Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: The developmental of causal reasoning in toddlers and infants. *Developmental Psychology, 42,* 1103–1115.

Sobel, D. M., & Kirkham, N. Z. (2007). Interactions between causal and statistical learning. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 139–153). Oxford, England: Oxford University Press.

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.

Tenenbaum, J. B., & Griffiths, T. L. (2003). Theory-based causal inference. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in neural information processing systems* (Vol. 15, pp. 35–42). Cambridge, MA: MIT Press.

Vadillo, M. A., & Matute, H. (2007). Predictions and causal estimations are not supported by the same associative structure. *Quarterly Journal of Experimental Psychology, 60,* 433–477.

Vandorpe, S., & De Houwer, J. (2005). A comparison of forward blocking and reduced overshadowing in human causal learning. *Psychonomic Bulletin & Review, 12,* 945–949.

Vandorpe, S., De Houwer, J., & Beckers, T. (2005). Further evidence for the role of inferential reasoning in forward blocking. *Memory & Cognition, 33,* 1047–1056.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of non-presentation of compound stimulus elements. *Learning and Motivation, 25,* 127–151.

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 53–76.

Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review, 8,* 600–608.

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.

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

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

Received February 5, 2008
Revision received June 25, 2009
Accepted July 8, 2009 ■