Developmental Psychology
2004, Vol. 40, No. 2, 162–176

0012-1649/04/$12.00 DOI: 10.1037/0012-1649.40.2.162

Causal Learning Across Domains

Laura E. Schulz and Alison Gopnik
University of California, Berkeley

Five studies investigated (a) children's ability to use the dependent and independent probabilities of events to make causal inferences and (b) the interaction between such inferences and domain-specific knowledge. In Experiment 1, preschoolers used patterns of dependence and independence to make accurate causal inferences in the domains of biology and psychology. Experiment 2 replicated the results in the domain of biology with a more complex pattern of conditional dependencies. In Experiment 3, children used evidence about patterns of dependence and independence to craft novel interventions across domains. In Experiments 4 and 5, children's sensitivity to patterns of dependence was pitted against their domain-specific knowledge. Children used conditional probabilities to make accurate causal inferences even when asked to violate domain boundaries.

The past two decades of research have demonstrated that young children understand cause and effect in a wide range of contexts. By the age of 4, children's folk physics includes knowledge about the causal relationship between object properties and object motion (Bullock, Gelman, & Baillargeon, 1982; Leslie & Keeble, 1987; Oakes & Cohen, 1990; Spelke, Breinlinger, Macomber, & Jacobson, 1992); their naive biology supports inferences about the causes of growth, inheritance, and illness (Gelman & Wellman, 1991; Inagaki & Hatano, 1993; Kalish, 1996); and their naive psychology allows them to explain the causes of human behavior in terms of emotions, desires, and beliefs (Flavell, Green, & Flavell, 1995; Gopnik & Wellman, 1994; Perner, 1991). In each of these domains, children can make appropriate predictions, provide causal explanations, and even make counterfactual claims (Harris, German, & Mills, 1996; Sobel, 2001; Wellman, Hickling, & Schult, 1997).

However, relatively little is known about how children acquire causal knowledge. Although researchers have abundant evidence of children's causal knowledge and have even traced changes in children's causal knowledge over development (see, e.g., Bartsch & Wellman, 1995; Gopnik & Meltzoff, 1997), there has been relatively little work experimentally investigating causal learning mechanisms. We know that children's causal knowledge changes, but we do not know how it changes. One way to answer the "how" question is to experimentally manipulate the kinds of evidence that children receive and then observe what causal inferences children draw on the basis of that evidence.

The few earlier studies that explored causal learning experimentally (in particular, Bullock et al., 1982, and Shultz, 1982) suggest that children can infer causal relationships using substantive knowledge about particular domains. In these studies, children were able to infer, for instance, that physical objects had to contact other objects to set them in motion and that physical effects required the transmission of force. Children appear to be able to apply domain-specific substantive principles about everyday physics to make new causal inferences about physical causal relations.

However, children might also learn causal relations in a more domain-general way by using what one might call formal principles of causal learning. In particular, children might infer causal relations from the pattern of dependence and independence among events. For example, children might observe whether the occurrence of one event makes the occurrence of another event more likely and then draw causal conclusions on the basis of that information. There is extensive evidence that adults make such inferences (Cheng, 1997, 2000; Shanks, 1985; Shanks & Dickinson, 1987; Spellman, 1996), but of course, adults have considerable experience and often explicit training in causal inference. Less is known about whether children use patterns of dependence to make causal inferences. There is some evidence that young children are able to use such principles in a particular task within the domain of physical causality (Gopnik, Sobel, Schulz, & Glymour, 2001). However, we do not know whether such formal learning mechanisms are specific to certain domains, whether they extend to learning causal relations in general, or whether they might even allow children to override domain-specific information.

The literature suggests that children do, in fact, have causal knowledge in a variety of domains, but this does not necessarily mean that this knowledge is the result of common learning mechanisms. Naive physics, naive psychology, and naive biology each represent distinct ways of knowing about the world. It seems possible that children might use one learning mechanism to understand mother's disappearance behind an occluder, another to understand why mother is happy, and yet another to understand why mother is coughing. Very different learning mechanisms might underlie children's understanding of very different events.

This research was supported by National Science Foundation Graduate Research Fellowship Award 99–136 to Laura E. Schulz and National Science Foundation Grant DLS0132487 to Alison Gopnik.

We are grateful to Catherine Griffith, Christine Schaeffer, Dan Slobin, Lori Markson, Dave Sobel, and Tamar Kushnir for helpful comments and suggestions. Thanks also go to the children and staff at Fruit and Flower, the Northeast Community Child Development Center, the Helen Gordon Child Development Center, the Goddard School, and the Berkeley Child Study Center.

Correspondence concerning this article should be addressed to Laura E. Schulz or Alison Gopnik, Department of Psychology, University of California, 3210 Tolman Hall #1650, Berkeley, CA 94720-1650. E-mail: laurasch@socrates.berkeley.edu or gopnik@socrates.berkeley.edu

In fact, many researchers have suggested that children's early knowledge of physics, psychology, and biology might originate in domain-specific modules (Leslie & Keeble, 1987) or from innate concepts in core domains (Keil, 1995; Spelke et al., 1992). These researchers have suggested that the human brain organizes information according to principles unique to each system of knowledge, much as Chomsky (1981) proposed for language learning. It is possible, therefore, that causal knowledge is accurate not because of general mechanisms designed to infer causal structure from evidence but because of specialized core systems that evolved for information-processing tasks particular to each domain (Leslie, 1994). A different but related view suggests that children might indeed use evidence to understand causal relationships, but only to extend and enrich core, domain-specific principles, not as a general technique for inferring new causal structure (Carey & Spelke, 1994).

Alternatively, common mechanisms of causal learning might apply across many domains. In principle, young children might use the pattern of dependence among events to infer the cause of a physical event, the cause of an emotion, or the cause of an illness. Formal inductive principles might be relevant to any causal learning task.

The particular formal inductive principle we examine in this article is the ability to *screen off* (Reichenbach, 1956) spurious associations and infer genuine causal relationships. If, for example, you notice a correlation between drinking wine and being unable to sleep, you might conclude that wine drinking causes insomnia. The problem is that the correlation between wine drinking and insomnia might be due to other events, also causally related to insomnia. For instance, if you usually drink wine at parties, parties might be a common cause of both wine drinking and insomnia, and wine and sleeplessness would be spuriously associated. These possibilities can be represented with simple graphs:

P (Parties) → W (Wine) → I (Insomnia)

(Parties cause wine drinking, which causes insomnia)

W (Wine) ← P (Parties) → I (Insomnia)

(Parties cause wine drinking and also cause insomnia).

Fortunately, it is possible to distinguish these graphs by observing the patterns of conditional dependence and independence among these events. If drinking wine while abstaining from parties is correlated with insomnia, but going to parties and abstaining from drink is not, then the first graph is correct. Formally, if W, I, and P are all dependent, and if W and I are still dependent in the absence of P but P is independent of I in the absence of W, then W screens off P from I. In our example, wine screens off partying as a causal explanation for insomnia.

Screening off can be understood as a special case of more general causal inference procedures. In recent years, researchers in artificial intelligence, statistics, and computer science have developed a set of formal procedures called causal Bayes nets, or causal directed graphical models (Glymour & Cooper, 1999; Pearl, 1988, 2000; Spirtes, Glymour, & Scheines, 1993), that provide algorithms for inferring the structure of a causal graph from patterns of dependency. Causal Bayes nets make a few general assumptions (in particular, the causal Markov and faithfulness assumptions; see Gopnik et al., 2004, for details) that specify which patterns of conditional probabilities will be generated by a particular causal structure. The algorithms provided by causal Bayes nets can infer the structure of the two simple graphs above, but they can also derive much more complex graphs.

Bayes net procedures infer causal structure from patterns of association, but they do not reduce causation to association or redefine causation in terms of associations. In particular, one arguably criterial feature of causal inference is that causal knowledge supports interventions. If we know two events are associated, we may predict that one event will typically follow the other. However, if we go beyond just the association and believe that there is a causal relation between the two events, we will also infer that acting to bring about the first event will change the probability that the second event will occur—an inference that will not hold for associated events that are not causally connected. Although a full discussion of an interventionist account of causal knowledge is beyond the scope of this article (see, e.g., Pearl, 2000, and Woodward, 2002, for more details), one unique feature of the Bayes net formalism is that it enables predictions about the effects of interventions (for a more complete account, see Glymour, 2001; Gopnik et al., 2004; Pearl, 2000).

For example, in both graphs above, wine is correlated with insomnia, and we can predict that the fact that we are drinking wine increases the probability that we will have insomnia. However, as the graph structures make clear, the probabilities exist for two different reasons. In the first graph, the two events are related because wine drinking causes insomnia, but in the second graph, they are related because the fact that we are drinking wine increases the probability that we are at a party, and parties cause insomnia. This leads to different predictions about interventions. If we want to avoid insomnia, we should stop drinking in the first case but should avoid partying in the second case. Thus, although having the right causal graph may not be essential for predicting an association between wine and insomnia, it is essential for predicting the effects of interventions that deliberately manipulate events. This is in contrast to the situation with associative learning models, which calculate the associative strength between pairs of variables but leave the steps linking associative strength, causal structure, and intervention external to the model.

Bayes net models can also take specific kinds of substantive prior causal knowledge into account in making new causal inferences. For example, we may know from independent sources that a causal link between wine and insomnia is more or less likely than one between partying and insomnia. Bayes net models can take that knowledge into account by adjusting the significance levels for determining conditional dependence and independence or by adjusting the prior probabilities of particular causal graphs. Thus background knowledge can be combined with dependency information in a precise way (for details, see Gopnik et al., 2004; Sobel, Tenenbaum, & Gopnik, in press; Tenenbaum & Griffiths, in press).

Previous research suggests that young children are indeed able to use screening off to make accurate and genuinely causal inferences in the domain of physical causality (Gopnik et al., 2001). In that research, children were shown a *blicket detector*, a box that lit up and played music when some objects, but not others, were placed on it. Children used screening off both to correctly infer which objects had the causal power to make the machine go and to design novel and appropriate interventions.

In the present study, we investigated the extent of children's use of formal inductive principles such as screening off and the relation between such formal principles and domain-specific knowledge. The experiments were designed to answer three questions: (a) Would children make formal inferences in domains other than physical causality, and how would those inferences compare with each other and with inferences in the physical domain (Experiments 1–3)? (b) Do children genuinely make causal inferences from conditional dependencies or might their behavior be explained by simpler strategies (Experiments 2–3)? (c) How do formal inference mechanisms interact with domain-specific knowledge (Experiments 4–5)? In Experiment 1, we extended the earlier blicket detector paradigm to explore whether children would use screening off to correctly identify the causes of biological and psychological events.

Experiment 1

Method

Participants

Thirty-two children ranging in age from 3 years 6 months to 5 years 2 months (mean age = 4 years 4 months) were recruited from urban area preschools. Approximately equal numbers of boys and girls participated. Sixteen children were randomly assigned to a test condition, and 16 children to a control condition. Although most children were from White, middle-class backgrounds, a range of ethnicities reflecting the diversity of the population was represented.

Materials

Biology. A monkey hand puppet, a glass vase, and two paper flowers were used for the biological screening-off task. The flowers were identical except for color. Each was affixed to a 20-cm-long bamboo stick.

Psychology. A stuffed-animal bunny, a wicker basket, and two small plastic animals (approximately 2.5 cm × 2.5 cm × 1.5 cm) were used for the psychological screening-off task. The plastic animals were a moose and an elephant.

Procedure

A female experimenter who was familiar to the children tested all the participants. Children were brought into a private game room in their school and sat facing the experimenter at a table. All children first participated in an unrelated experiment. The order of domain presentation was counterbalanced across participants.

Biology. All children were introduced to the vase and the monkey puppet. Children were told, "Monkey likes to smell flowers, but some flowers make Monkey sneeze. Will you help me figure out which flowers Monkey likes to smell and which flowers make Monkey sneeze?" All trials began when the monkey came up to smell the flower. The effect either occurred or failed to occur, and the trials were terminated by moving the monkey away from the vase.

Children in the test condition were shown two flowers, A and B. The experimenter placed each flower in the vase by itself, in counterbalanced order, and brought the monkey puppet up to smell each flower. Flower A made the monkey sneeze; Flower B did not. The experimenter then placed both flowers in the vase together and brought the monkey puppet up to smell both flowers at once. The monkey sneezed. The experimenter brought the monkey up to smell both flowers together a second time, and again the monkey sneezed. The experimenter then removed both flowers from the vase and asked the child to "give me the flower that makes Monkey sneeze." If the children are using formal screening-off principles to make causal judgments, Flower A should screen off Flower B from the effect and children should give the experimenter Flower A. Flower A and sneezing were dependent even in the absence of Flower B, whereas Flower B and sneezing were independent in the absence of Flower A.

Children in the control condition were also shown two flowers, A and B. The experimenter placed each flower in the vase by itself, in counterbalanced order, and then brought the monkey puppet up to smell each flower three times. Flower A made the monkey sneeze all three times; Flower B did not make the monkey sneeze the first time, but it did make the monkey sneeze the next two times. The experimenter then removed both flowers from the vase and asked the child to "give me the flower that makes Monkey sneeze."

Note that the contingencies between the flowers and the sneezing were identical in the test and control conditions. Flower A was associated with the effect 100% of the time, and Flower B was associated with the effect only 66% of the time. However, in the control condition, the flowers were never presented simultaneously, so screening off was not an option. Because each flower makes the monkey sneeze independent of the other flower, the children should choose between the flowers at chance levels.

Psychology. Except for features relevant to the domain, this protocol was identical to the biology protocol. All children were introduced to the basket and the bunny. Children were told, "Bunny likes some animals, but some animals scare him. Will you help me figure out which animals Bunny likes and which animals make Bunny scared?" All trials began when the bunny came up to look in the basket. The effect either occurred or failed to occur, and the trials were terminated by moving the bunny away from the basket.

Children in the test condition were shown plastic animals, A and B. The experimenter placed each animal in the basket by itself, in counterbalanced order, and brought the bunny up to look at each animal. To demonstrate the bunny being scared, the experimenter had it say "Eek!" and back away. When the bunny was not scared, the experimenter briefly placed the bunny in the basket with the animal. In the test condition, Animal A screened off Animal B from the effect; in the control condition, the animals were always presented independently. In both conditions, Animal A was associated with the effect 100% of the time and Animal B was associated with the effect 66% of the time. The experimenter then removed both animals from the basket and asked the child to "give me the animal that makes Bunny scared."

Results

Preliminary analyses revealed no effect of age or order of domain presentation on the children's responses for either the test or control conditions. Children's responses to the test and control conditions across the two domains are presented in Table 1.

Biology

In the test condition, 100% of the children chose Flower A as the causal object. By contrast, in the control condition, children chose

Table 1
Number of Children per Type of Response in Experiment 1

	Biology		Psychology	
Response	Test	Control	Test	Control
Chose Object A (the 100% object)	16 (100)	9 (56)	14 (88)	9 (56)
Chose Object B (the 66% object)	0 (0)	7 (44)	2 (12)	7 (44)

Note. Numbers in parentheses are percentages. n = 16 per condition.

between the flowers at chance, $\chi^2(1, n = 16) = 0.25$, *ns*, 56% of the children choosing Flower A and 44% of the children choosing Flower B. Children were significantly more likely to choose Object A as the causal object in the test condition than in the control condition, $\chi^2(1, N = 32) = 8.96$, $p < .01$.

Psychology

In the test condition, 87.5% of the children chose Animal A as the causal object, and 12.5% of the children chose Animal B. Children in the test condition chose the causal animal significantly more often than would be expected by chance, $\chi^2(1, n = 16) = 9.00$, $p < .01$. By contrast, in the control condition, children chose between the animals at chance, $\chi^2(1, n = 16) = 0.25$, *ns*, 56% of the children choosing Animal A and 44% of the children choosing Animal B. Again, children were significantly more likely to choose Object A as the causal object in the test condition than in the control condition, $\chi^2(1, N = 32) = 3.86$, $p < .05$.

Discussion

These results are consistent with the inferences that would be made in the Bayes net formalism. Suppose A and B represent the state of Objects A and B (present or not) and that C represents the state of the monkey or the bunny (sneezing or not; scared or not). Assuming that the observed frequencies are representative of the underlying probabilities, and that the causes are generative and noninteractive, the state of C is independent of the presence or absence of Object B, conditional on the presence of Object A.

Intuitive notions of dependence, independence, and conditional dependence and independence can be translated into exact statements about probabilities. Formally, if B and C are independent conditional on A, $Pr(B, C \mid A) = Pr(B \mid A) * Pr(C \mid A)$. However, the state of C is not independent of the presence or absence of Object A: $Pr(A, C \mid B) \neq Pr(A \mid B) * Pr(C \mid B)$. Conditioning on the value of B does not make A and C independent.

Applied to this case, a Bayes net learning algorithm will construct the model depicted in Figure 1.[1] This graph represents all the possible causal graphs that are consistent with the Markov and faithfulness assumptions and the patterns of conditional dependence we just described.

The graph says that A causes C and that B does not. (It also says that there is some undetermined, and perhaps unobserved, causal link between A and B, represented by the circles and the ends of the edge connecting those variables. In fact, there is such a link—namely, the experimenter, who ensures that the two objects are most often present at the same time.)

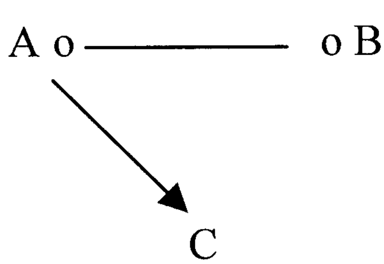


Figure 1. Graph representing inference that A screens off B as a cause of C.

The fact that children also believe, across domains, that A causes C and that B does not suggests that preschool children can use patterns of dependence and independence to screen off one variable in favor of another, even after only a single exposure to the relevant independent and dependent probabilities, and that they can do so in the biological and psychological domains as well as the physical domain.

Experiment 2

Although the children in Experiment 1 were presented with evidence about conditional dependence and independence, it is possible that they may have ignored part of the evidence. Specifically, the children may have attended only to the events involving a single candidate cause. They may have observed the dependence between A and the effect, and the independence between B and the effect, while ignoring the dependence between A and B *together* and the effect. In this case, the children could have used the simple pattern of dependence and independence to choose Flower A, rather than attending to the conditional probabilities of events. A similar criticism might apply to the original physical screening-off task in Gopnik et al. (2001).

Some previous research with the blicket detector (Gopnik et al., 2004; Sobel et al., in press) weighs against this interpretation; findings in the domain of physical causality suggested that children do attend to trials involving more than one object. However, those findings involved only physical causality, and in those experiments there were still some trials in which the causal object activated the detector by itself. If preschool children are genuinely sensitive to conditional probabilities, then they should also be able to infer causal relationships even if they are never presented with any evidence about a single candidate cause (i.e., if they never see either A or B by itself). In Experiment 2, we created a new version of the biology screening-off task to test whether children could interpret such complex patterns of evidence.

Children were presented with four candidate causes: A, B, C, and a distractor, D. They were given evidence that A and C together produced the effect, that B and C together produced the effect, and that A and B together failed to produce the effect. They were then given A, B, C, and D and asked to choose the causal object. Would preschool children be able to use conditional probabilities to make this more complex normative causal inference?

Method

Participants

Twenty-eight children ranging in age from 3 years 9 months to 5 years 4 months (mean age = 4 years 7 months) were recruited from an urban area preschool. Approximately equal numbers of boys and girls participated. The children were randomly assigned to a test condition and a control condition. Although most children were from White, middle-class backgrounds, a range of ethnicities reflecting the diversity of the population was represented.

[1] This assumes that the experimenter's role in setting the state of A and B eliminates the possibility of a common cause of A and C and of B and C. It also assumes, as noted above, that the children take the frequencies as representative of the probabilities, despite the small sample size.

Materials

The same materials used for the biological task in Experiment 1 were used in Experiment 2 except that four rather than two paper flowers were used. Each flower was a different color.

Procedure

A female experimenter who was familiar to the children tested all the participants. Children were brought into a private game room in their school and sat facing the experimenter at a table. Children first participated in an unrelated experiment.

All children were introduced to the monkey puppet. An empty vase was placed on the table and children were told, “Monkey likes to smell flowers, but some flowers make Monkey sneeze. Let’s see what happens.” In each condition, children were shown Flowers A, B, C, and D and were asked to name the color of each flower. The flowers were presented in random order. Trials began when the puppet came up to smell the flowers in the vase, and trials were terminated by moving the puppet away from the vase.

Test condition. The experimenter removed Flowers A, B, and C from view, leaving Flower D on the table in front of the child. The experimenter then placed Flowers A and C in the vase. She said, “Let’s try the (red) flower and the (blue) flower.” Monkey came up to smell Flowers A and C and sneezed. Monkey backed away from the vase, came back up to smell the flowers, and sneezed again. The experimenter removed Flowers A and C and replaced them with Flowers B and C. She said, “Now let’s try the (blue) flower and the (yellow) flower.” Monkey came up to smell the flowers twice, and each time the monkey sneezed.

The experimenter removed Flowers B and C and replaced them with Flowers A and B. She said, “Now let’s try the (red) flower and the (yellow) flower.” Monkey came up to smell the flowers a single time, and this time Monkey did not sneeze. The experimenter removed Flowers A and B from the vase, picked up all four flowers, and arranged them in random order on the table in front of the child. She asked the child, “Can you give me the flower that makes Monkey sneeze?”

Control condition. The experimenter removed Flowers A, B, and C from view, leaving Flower D on the table in front of the child. The experimenter then placed Flower C in the vase. She said, “Let’s try the (red) flower.” Bunny came up to smell Flower C and sneezed. Bunny backed away from the vase, came back up to smell the flower, and sneezed again. This was repeated so that Bunny sneezed a total of four times in response to Flower C.

The experimenter removed Flower C and placed Flower A in the vase. She said, “Now let’s try the (blue) flower.” Bunny came up to smell Flower A and sneezed. Bunny backed away, smelled Flower A again, and sneezed. Bunny backed away, smelled Flower A a third time, and did not sneeze. Flower A was removed, and the experimenter repeated the same protocol for Flower B.

The experimenter then picked up all four flowers and arranged them in random order on the table in front of the child. She asked the child, “Can you give me the flower that makes Bunny sneeze?”

Note that the flowers and sneezing are associated with identical frequencies in the test and control conditions: Flower C produces the effect four out of four times, whereas Flowers A and B each produce the effect two out of three times. If children are using frequency information to make causal judgments, then children should choose Flower C in both the test and control conditions. However, if children are using the conditional probabilities of events to make causal inferences, then the children should choose Flower C in the test condition but should choose at chance between the three flowers in the control condition because all three flowers in the control condition independently produce the effect. Finally, if children in the test condition are only attending to the last trial (i.e., the trial in which A and B both fail) and not to the positive trials involving C, then they should choose at chance between C and the distractor.

Results and Discussion

In the test condition, children chose the causal flower both significantly more often than would be expected by chance, $\chi^2(3, n = 14) = 21.43$, $p < .01$, and significantly more often than they chose all other flowers, $\chi^2(1, n = 14) = 4.57$, $p < .05$. Seventy-nine percent of the children chose the causal flower (Flower C), 14% chose Flower A, 7% chose Flower B, and no children chose the distractor. By contrast, in the control condition, children chose between the flowers at chance, $\chi^2(3, n = 14) = 1.42$, *ns*. Twenty-one percent of the children chose Flower C (the 100% flower), 29% chose Flower A, 36% chose Flower B, and 14% chose the distractor. Although Flower C was associated with the effect four out of four times in both the test and control conditions, children were significantly more likely to screen off and choose Flower C in the test condition than they were to choose Flower C in the control condition, $\chi^2(1, N = 28) = 9.14$, $p < .01$.

In order to infer that Flower C screened off Flowers A and B from the effect, children in the test condition had to keep track of a complex pattern of conditional dependence and independence. Monkey sneezing to Flowers A and C, and sneezing to Flowers B and C, provided evidence that sneezing was independent of the presence or absence of Flower A, conditional on the presence of Flower C. (If C was in the vase, Monkey sneezed whether A was there or not.) The same data provided evidence that, conditional on the presence of C, sneezing was independent of the presence or absence of Flower B.

By contrast, sneezing to Flowers A and C combined with the lack of sneezing to Flowers A and B provided evidence that conditioning on the presence of Flower A did not make sneezing independent of the presence or absence of Flower C. (If A was always in the vase, Monkey sneezed when C was there and didn’t sneeze when C was not.) Similarly, sneezing to B and C combined with the absence of sneezing to A and B demonstrated that conditioning on the presence of B did not make sneezing independent of the presence or absence of Flower C. Because conditioning on A did not make C and the effect independent, and conditioning on B did not make C and the effect independent, children could infer that C and the effect were unconditionally dependent.

Only having tracked this pattern of conditional dependence and independence could children engage in screening off. Because A and B were independent of sneezing conditional on C, but C and sneezing were unconditionally dependent, children could infer that Flower C screened off A and B from the effect. Four-year-old children were able to keep track of the conditional probabilities and make accurate causal inferences, even in the biological rather than the physical domain and even given this relatively complex pattern of evidence.

Experiment 3

The inferences required in Experiment 2 were more sophisticated than those required in Experiment 1; however, the screening-off reasoning described in both studies bears some resemblance to the phenomenon of *blocking* in classical conditioning. In blocking, an animal that receives a shock accompanied by a light and then receives the shock accompanied by a light and a tone shows a fear response to the light but not the tone. It is possible that the children might simply have associated the sneezing (or fear) with one flower (or animal) rather than the other instead of making a genuinely causal inference. They might not have understood the test questions as causal questions at all but may simply have answered with the object that was most strongly associated with the relevant effect.

However, if children can use screening off to make genuinely causal inferences, then children ought to be able to do two things that go beyond simple association. First, they should be able to anticipate the effect of novel interventions. Second, they should be able to combine formal principles such as screening off with what we might call substantive causal principles: commonsense prior knowledge about causal relations.

As noted above, Bayes net learning algorithms would infer the causal structure in Figure 1 from the patterns of causal dependence in Experiment 1. This structure in turn generates predictions about interventions. In particular, it implies that an intervention on A will change the value of C but that an intervention on B will not have this effect. If children can combine such formal information about interventions with specific substantive knowledge about *how* to perform interventions, then children should be able to produce interventions that are substantively novel but formally normative. A likely substantive causal principle is that if a cause produces an effect, removing the cause will remove the effect. This is not a necessary principle, of course, but it is a plausible pragmatic assumption about many types of causal relations in many domains. If, for instance, children really think that red flowers have the causal power to make Monkey sneeze, then they could infer on formal grounds that intervening on the red flowers will influence Monkey's sneezing. If they also know that removing a cause tends to remove its effect, they should infer that removing the red flowers might make Monkey stop sneezing. The formal prediction about interventions tells them which object to intervene on, and their prior substantive knowledge tells them how to intervene on that object.

On the other hand, if children are merely associating the red flowers and the sneezing, along the lines of classical conditioning, then children should not draw this inference, nor should they be able to craft an appropriate intervention. Although the principles of classical conditioning predict that animals in a blocking experiment can learn to associate a shock with a light instead of a tone, such principles do not predict that animals will act to extinguish the light. Except for species-specific defense reactions (freezing, attack, or flight; see Bolles, 1970), animals in classical conditioning paradigms do not spontaneously intervene to prevent effects by removing causes. Of course, animals can learn through trial and error to make appropriate interventions (i.e., to push a lever in order to avoid a shock; see, e.g., Sidman, 1953); however, such learning depends on reinforcement and operant conditioning. Animals can predict the effects of an intervention once they have performed it themselves, but they do not appear to use information about the dependencies among events to infer causal structure and to design interventions that they have never performed before (see Gopnik et al., 2004, for further discussion).[2]

Previous research suggested that children are able to generate such interventions in the domain of physical causality (Gopnik et al., 2001). In Experiment 3 we extended these findings to the domains of biology and psychology and also eliminated an alternative explanation of the earlier intervention findings.

In the Gopnik et al. (2001) study, Block A was placed on a machine and nothing happened. Block A was removed, and Block B was placed on the machine. The machine activated, and A was replaced on the machine next to B. Children were asked to stop the machine, and they did so by removing B but not A. This task involved a novel action, but one could argue that the action, rather than being truly novel, was designed to recreate a state of affairs that the child had already seen (the machine failing to activate when only A was present). By contrast, in Experiment 3, children were given three objects: one causal and two noncausal. The children were shown each object individually and all three objects together, but they were never shown the two noncausal objects together without the causal object. Thus, for children to prevent the effect by removing only the causal object, they had to both produce a new intervention (stopping the machine) and generate a new state of affairs (in which the two noncausal objects were present and the causal object was absent).[3]

Experiment 3 also enabled us to address a potential concern about Experiments 1 and 2. In those experiments, children were asked to identify a single causal object in a forced-choice design (they were asked to show the experimenter the object that made the monkey sneeze or the bunny scared). Although the children were able to screen off and accurately identify the causal object, it is possible that they believed that multiple objects had causal properties, with the chosen object simply being more causal than the others. By allowing the children to craft an open-ended intervention, we could assess the specificity of children's causal judgments.

Method

Participants

Thirty-six children ranging in age from 3 years 8 months to 5 years 5 months (mean age = 4 years 8 months) were recruited from urban area preschools. Eighteen children were randomly assigned to a test condition, and 18 to a control condition. Approximately equal numbers of boys and girls participated. Although most children were from White, middle-class backgrounds, a range of ethnicities reflecting the diversity of the population was represented.

[2] This is not to say that all animal interventions can be explained by the principles of operant conditioning. Indeed, animals' ability to maintain a learned avoidance response has been one of the bugaboos of behaviorist accounts of learning. Although operant conditioning can explain how an animal learns to press a lever to avoid a shock, associationist accounts cannot easily explain why the behavior is maintained rather than extinguished after repeated instances of successful avoidance (see, e.g., Schwartz, 1978, for discussion). In response, some researchers (Seligman & Johnston, 1973) have proposed cognitive accounts of avoidance learning in animals; that is, they have suggested that animals are also making genuinely causal inferences rather than simply making associations. Thus, the distinction between associationist and causal accounts of learning should not necessarily be conflated with a distinction between animal and human cognition. The point here is that children's inferences go well beyond those of classical or operant conditioning; it is possible that animals' inferences also go beyond these procedures, but confirmation of this awaits further research.

[3] We are indebted to an anonymous reviewer for suggestions that contributed to the design in Experiment 3.

Materials

The same materials used in Experiment 1 were used in Experiment 3 except that a third flower was added to the biological screening-off task and a third animal (a polar bear) was added to the psychological screening-off task.

Procedure

A female experimenter who was familiar to the children tested all the participants. Children were brought into a private game room in their school and sat facing the experimenter at a table. Children first participated in an unrelated experiment. The order of domain presentation (biological or psychological) was counterbalanced across participants.

Biology. All children were introduced to the monkey puppet and told, "Monkey likes to smell flowers, but some flowers make Monkey sneeze. Let's see what happens." The children were shown Flowers A, B, and C and asked to name the color of each flower. The experimenter then placed the vase on the table. Trials began when the monkey came up to smell the flowers in the vase, and trials were terminated by moving the monkey away from the vase.

Children in the test condition were asked to place Flower A in the vase (flower color was counterbalanced across participants). Monkey came up to smell Flower A and did not sneeze. Children were asked to remove Flower A and replace it with Flower B. Again Monkey came up to smell the flower and did not sneeze. Then children were asked to place Flower C in the vase. This time Monkey sneezed. While Monkey was sneezing, the experimenter added Flowers A and B back to the vase. Monkey continued to sneeze. The experimenter then moved Monkey away from the vase and asked the child, "Can you make it so Monkey won't sneeze?"

If children use screening-off assumptions to draw causal conclusions and are able to combine those assumptions with the substantive knowledge that removing causes often removes effects, they should remove Flower C rather than Flowers A or B. Children saw that Flower C and the monkey sneezing were dependent even in the absence of Flowers A and B but that Flowers A and B were independent of the effect in the absence of Flower C. This should lead the children to conclude that C, not A or B, caused the effect and that they should remove C, but not A or B, to stop the effect. Note also that because the children never saw Flowers A and B together in the vase without C, children could not merely associate the combination of A and B with not sneezing. Rather, the children in the test condition had to generate both a novel action and a novel state of affairs.

Children in the control condition saw an identical set of events except that this time all three flowers, A, B, and C, independently made the monkey sneeze. Monkey came up to smell Flower A and sneezed. The experimenter moved Monkey away from the vase. Children were asked to remove Flower A and replace it with Flower B. Monkey came up to Flower B and sneezed. Again the experimenter moved Monkey away. Then children were asked to place Flower C in the vase. Monkey sneezed. While Monkey was sneezing, the experimenter added Flowers A and B back to the vase. Monkey continued to sneeze. The experimenter then moved Monkey away from the vase and asked the child, "Can you make it so Monkey won't sneeze?"

This time, if children understood the causal properties of the flowers and the substantive causal principle that removing causes often removes effects, they should have removed all three flowers from the vase. In this case, children saw that each flower made the monkey sneeze independent of any other flower. Children should have concluded that all three flowers caused the effect and that all three flowers would have to be removed to stop the effect.

Psychology. This protocol was formally identical to the biology protocol except for features relevant to the domain. Trials began when the bunny came up to look in the basket and were terminated by moving the bunny away from the basket, and the bunny either acted scared or did not act scared. Children were asked to "Make it so Bunny won't be scared."

Children in the control condition saw a set of events identical to the ones in the control condition in the biology task except that this time all three animals, A, B, and C, independently made Bunny scared. This time, if children understood the causal properties of the animals and the substantive causal principle that removing causes often removes effects, they should have removed all three animals from the basket.

Results and Discussion

Preliminary analyses revealed no effect of order of domain presentation on the children's responses. Children's responses in the test and control conditions across the two domains are presented in Table 2.

Biology

Children could make a variety of responses to the "Make it so Monkey won't sneeze" question. They could remove Flower C, they could remove all the flowers, they could remove any other single flower or combination of two flowers, or they could make another response altogether (i.e., removing the entire vase).

Children were significantly more likely to screen off and remove only Flower C in the test condition than in the control condition, $\chi^2(1, N = 36) = 16.46$, $p < .01$. Children were also significantly more likely to remove all three flowers in the control condition than in the test condition, $\chi^2(1, N = 36) = 18.78$, $p < .01$.

Within the test condition, 14 of the 18 children (78%) screened off and removed only Flower C, whereas only 2 children (11%) removed all the flowers. Children in the test condition were significantly more likely to screen off than to remove all the flowers, $\chi^2(1, n = 16) = 9.00$, $p < .01$. Within the control condition, only 1 child (5%) removed only Flower C, whereas 16 of the 18 children (89%) removed all of the flowers. Children in the control condition were significantly more likely to remove all the flowers than to remove only Flower C, $\chi^2(1, n = 17) = 13.23$, $p < .01$.

Psychology

Children could make a similar variety of responses to the psychological question. Children were significantly more likely to screen off and remove only Animal C in the test condition than in the control condition, $\chi^2(1, N = 36) = 12.04$, $p < .01$. Children were also significantly more likely to remove all three animals in the control condition than in the test condition, $\chi^2(1, N = 36) = 9.11$, $p < .01$.

Table 2
Number of Children per Type of Response in Experiment 3

Response	Biology		Psychology	
	Test	Control	Test	Control
Removed C	14 (78)	1 (5)	12 (67)	1 (5)
Removed all	2 (11)	16 (89)	5 (28)	15 (83)
Other	2 (11)	1 (5)	1 (5)	2 (11)

Note. Numbers in parentheses are percentages (due to rounding, percentages may not sum to 100). $n = 18$ per condition.

Within the test condition, 12 of the 18 children (67%) screened off and removed only Animal C, whereas 5 children (28%) removed all the animals. Children in the test condition were marginally more likely to screen off than to remove all the animals, $\chi^2(1, n = 17) = 2.88, p < .10$. Within the control condition, only 1 of the children (5%) removed only Animal C, whereas 15 of the 18 children (83%) removed all of the animals. Children in the control condition were significantly more likely to remove all the animals than to remove only Animal C, $\chi^2(1, n = 16) = 12.25$, $p < .01$.

Finally, we examined children's performance across the two domains. Children were remarkably consistent across domains. There was no significant difference between domains in children's tendency to remove the causal object rather than anything else in the test conditions, McNemar's $\chi^2(1, n = 18) = 0.50$, *ns*, or to remove all the objects rather than anything else in the control conditions, McNemar's $\chi^2(1, n = 18) = 0.00$, *ns*.

Recall that in this experiment, the children in the test condition had an opportunity to associate Monkey not sneezing and Bunny not being scared with the presence of Object A by itself, the presence of Object B by itself, and the absence of all three objects. However, when asked to inhibit the effect, the majority of children created something they had never had an opportunity to observe: A and B together without C. The combination of the novel intervention and the novel result suggests that preschool children's ability to make formal causal inferences extends well beyond mechanisms such as blocking in classical conditioning and extends to the biological and psychological domains as well as the physical domain.

Experiment 4

The results of Experiments 1, 2, and 3 suggest that children's ability to infer causal relationships from patterns of dependence and independence is domain independent. Children make very similar inferences in physical, biological, and psychological domains. However, it is possible that children use patterns of dependence to differentiate equally plausible causal candidates within a domain (i.e., the causal power of one flower vs. another) but that domain-specific knowledge restricts the range of evidence children are willing to consider in the first place. Inferences about how biological agents cause illness are typical of young children's biological reasoning, and inferences about how threats cause fear are typical of young children's psychology. One might argue that formal inference procedures enrich knowledge in each domain but do not override or change that knowledge or force children to cross domain boundaries. How broad is the scope of children's domain-general causal reasoning? What happens when domain-general causal learning procedures conflict with domain-specific knowledge?

Recent research has largely sought to demonstrate that children's causal reasoning respects domain-specific principles (see, e.g., Hickling & Wellman, 2001). To our knowledge, only a single study (Notaro, Gelman, & Zimmerman, 2002) has looked at children's ability to make causal judgments across domain boundaries. In that study, the researchers looked at causal judgments about psychogenic illness and found that preschool children were reluctant to attribute effects in one domain to causes from another. Specifically, preschoolers rarely endorsed psychological causes for biological effects (i.e., worrying as a cause of headaches).

Importantly, however, that study relied on children's prior causal knowledge and did not provide the children with any new evidence about psychogenic causes. All systems of causal inference are limited by the evidence available. If children's prior knowledge provided evidence only for the relationship of biological causes to biological effects, then even a domain-general method of causal inference would derive a domain-specific representation of causal structure. However, if children are able to build and revise causal inferences directly from evidence, rather than exclusively from prior knowledge or domain-specific modules, then domain-specific judgments ought to be defeasible. Given appropriate evidence, children ought to be able to override prior domain-specific inferences and reason about events that cross the boundaries of domains.

In Experiment 4, we replicated the test condition of Experiment 3, but we pitted children's domain-specific knowledge (e.g., that physical effects have physical causes) against evidence from patterns of dependence. To ensure that children did indeed initially assume that causal inferences would respect domain boundaries, we introduced a new baseline condition. In the baseline condition, children were given no dependence information and were simply asked to predict whether a domain-appropriate or domain-inappropriate candidate event would cause the test events. In the test condition, children were given dependence information that indicated that the domain-inappropriate candidate was actually the correct cause. If children rely primarily on domain-specific knowledge in their causal judgments, then children should choose the domain-appropriate causes in both the test and baseline conditions. If children can use evidence from patterns of dependence to override domain-specific knowledge, then children should choose the domain-appropriate causes in the baseline condition but should choose the domain-inappropriate causes in the test condition.

This design also enabled us to eliminate a possible confound in Experiments 1–3. Sneezing and fear seem to fall straightforwardly into the domains of biology and psychology, respectively, but we had no independent evidence that the children in our studies classified the phenomena in this way. It is conceivable that the children in Experiments 1–3 may have applied common methods of causal inference in part because they failed to categorize sneezing and fear as distinctively biological and psychological events, respectively. The baseline condition in this experiment tested whether children really did categorize the relevant events differently.

Method

Participants

Thirty-two children ranging in age from 3 years 8 months to 5 years 4 months (mean age = 4 years 6 months) were recruited from urban area preschools. Sixteen children were randomly assigned to the baseline condition, and 16 children were assigned to the test condition. One child assigned to the test condition chose not to participate and was replaced. Approximately equal numbers of boys and girls participated. Although most children were from White, middle-class backgrounds, a range of ethnicities reflecting the diversity of the population was represented.

Materials

Physical effects task. An Imagine Nation (Imagine Nation Group, Los Angeles, CA) blue and red plastic toy pottery wheel with foot pedal was used as a "noise-making machine" in the physical effects task. The pottery wheel was turned upside down throughout the experiment, and the foot pedal was concealed underneath a table. No child identified the machine as a pottery wheel or gave any indication of suspecting the existence of the foot pedal. Two 6-cm-diameter round magnets (one pink and blue, the other purple and yellow) that could be placed on top of the machine were used as the "buttons" in the physical effects task. In the baseline task, children also saw a picture of the experimenter talking to the machine.

Psychological effects task. Two drawings of silly faces were used in the psychological effects task. One face was drawn on a green index card, the other on a yellow index card. A 10 cm × 4 cm × 3 cm "fake" switch box was also used in this task. The box had an on/off switch and a 3-ft electrical cord; however, the cord did not attach to anything, and the switch was nonfunctional. The box was painted with yellow polka dots. The same vase used in Experiments 1 and 2 was also used in this task.

Baseline warm-up tasks. Three sets of materials were used as training tasks in the baseline condition. One set consisted of a lemon and two metallic hardware items (a pin and a ring). Another set consisted of scissors, a drawing compass, and a tennis ball. The third set consisted of a pencil and two different ballpoint pens. The wicker basket used in Experiments 1 and 2 was also used.

Procedure

A female experimenter who was familiar to the children tested all the participants. In the test condition, a confederate assisted the experimenter. Children were brought into a quiet room in their school and sat facing the experimenter at a table. Children first participated in an unrelated experiment.

Baseline condition. Children in the baseline condition were told, "We're going to play a sorting game" and were given three warm-up tasks. The warm-up tasks were introduced so that the children knew that they could put either the two similar objects, the one outlier object, or all three objects together in the basket. These tasks also served as a control to show that children would indeed produce all three types of responses. In one warm-up task, the experimenter placed the pin, the ring, and the lemon on the table and said, "Here are some things that might be fruit. Could you put the things that really are fruit into the basket?" In another warm-up task, the experimenter placed the scissors, the drawing compass, and the tennis ball on the table and said, "Here are some things that might be sharp. Could you put the things that really are sharp into the basket?" In a third warm-up task, the experimenter placed the pencil and two pens on the table and said, "Here are some things that you might write with. Could you put the things you really could write with into the basket?" The warm-up tasks were presented in random order. All children passed the warm-up tasks.

After the warm-up tasks, the children received a physical effects task and a psychological effects task (domain order was counterbalanced across participants). In the physical effects task, the experimenter placed the machine on the table in front of the child and said, "Here's a machine that can make noise. The machine's not plugged in right now, but here are some ways that I might be able to make the machine go." The experimenter gave the child a choice between a domain-inappropriate cause and two domain-appropriate causes. For the domain-inappropriate cause, the experimenter showed the child a picture of herself talking to the machine and said, "I might make the machine go by talking to the machine and saying, 'Machine, please go!'" For the domain-appropriate causes, the experimenter brought out each magnet button in turn, placed it on the machine, and said, "I might make the machine go by putting this button on the machine like this." This procedure was analogous to the physical screening-off task in Gopnik et al. (2001), in which blocks were placed on the blicket detector to make it go. The experimenter placed all three choices on the table in front of the child and said, "Can you put the ways you really could make the machine go into the basket?"

For the psychological effects task, the experimenter said, "Here are some things that might make a person giggle." For the domain-inappropriate cause, the experimenter brought out the switch and said, "Here's a switch." For the two appropriate causes, the experimenter brought out each drawing of a face and said, "Here's a silly face." The experimenter placed all three choices on the table in front of the child and then said, "Can you put the things that really could make a person giggle into the basket?" In both tasks, the order of presentation and the position of the three causes were counterbalanced across participants.

If the children have no domain-specific assumptions about what might cause machines to go or people to giggle, then they should choose at chance. If the children believe that physical causes are likely to produce physical effects and that psychological causes will produce psychological effects, then the children should select only the domain-appropriate causes.

Test condition. Children in the test condition received a physical effects task and a psychological effects task, with domain order counterbalanced across subjects. For the physical effects task, the machine was placed on the table and the children were told, "This is my machine. Some things make my machine make noise. Can you help me figure out what makes the machine go?"

Children were presented with two domain-appropriate causes (buttons) and one domain-inappropriate cause (talking to the machine). The experimenter placed Button A on the machine (the particular button was counterbalanced across subjects). Nothing happened. The experimenter removed Button A and placed Button B on the machine. Again the machine did nothing. The experimenter removed Button B, and then the experimenter said, "Machine, please go!" The experimenter surreptitiously pressed the foot pedal, and the machine began to make a loud whirring noise. While the machine was whirring, the experimenter placed Buttons A and B back on the machine. After 10 s, the experimenter passed the machine to the child and said, "This machine is making a lot of noise. Can you make the machine be quiet?"

Children saw that talking to the machine and the machine starting were dependent even in the absence of Buttons A and B but that Buttons A and B were independent of the effect in the absence of talking to the machine. If children are screening off, they should say something like "Machine, please stop!" Conversely, if the children are relying on domain-specific assumptions to guide their causal inferences and believe that physical causes (i.e., buttons) are more likely than psychological causes (i.e., talking) to produce physical effects, then the children should remove the buttons from the machine.

For the psychological effects task, the experimenter introduced the child to a confederate by saying "This is my friend Catherine. Catherine is pretty silly. She giggles a lot. Can you help me figure out what makes Catherine giggle? Catherine, close your eyes."

The psychological effects task was formally identical to the physical effects task. The two domain-appropriate causes were drawings of silly faces; the domain-inappropriate cause was a switch. Each silly face drawing was placed in the vase in turn; the confederate opened her eyes and failed to giggle. However, when the experimenter flipped the switch, the confederate immediately began giggling. While the confederate was giggling, the experimenter returned Silly Faces A and B back to the vase within the confederate's line of gaze. After 10 s, the experimenter said, "Catherine is giggling a lot. Can you make Catherine stop giggling?"

If children use screening-off assumptions to draw causal conclusions, children should flip the switch. Conversely, if the children rely on domain-specific assumptions, then the children should remove the drawings of silly faces from the confederate's sight.

Results and Discussion

Preliminary analyses revealed no effect of order of domain presentation on the children's responses. In both the baseline and test conditions, the children could choose the domain-inappropriate cause, choose one or both domain-appropriate causes, or make another response altogether (i.e., they could make no response or say "I don't know"). Children's responses are presented in Table 3.

Physical Effects Tasks

Children were significantly more likely to choose the domain-inappropriate object in the test condition than in the baseline condition, $\chi^2(1, N = 32) = 16.13$, $p < .01$. Children were also significantly more likely to choose the domain-appropriate objects in the baseline condition than in the test condition, $\chi^2(1, N = 32) = 28.12$, $p < .01$.

Within the test condition, children were more likely to choose the domain-inappropriate object than to choose the domain-appropriate objects or to make another response, $\chi^2(2, n = 16) = 14.09$, $p < .01$. Twelve of the 16 children (75%) screened off and chose the domain-inappropriate cause (i.e., they said "Stop," "Please stop," or "Machine, please stop"), whereas no child chose the domain-appropriate causes (the buttons). By contrast, in the baseline condition, no child chose the domain-inappropriate cause, and 100% of the children chose the domain-appropriate causes.

Psychological Effects Tasks

Children were significantly more likely to choose the domain-inappropriate object in the test condition than in the baseline condition, $\chi^2(1, N = 32) = 18.66$, $p < .01$. Children were also significantly more likely to choose the domain-appropriate objects in the baseline condition than in the test condition, $\chi^2(1, N = 32) = 28.12$, $p < .01$.

Within the test condition, children were more likely to choose the domain-inappropriate object than to choose the domain-appropriate objects or to make another response, $\chi^2(2, n = 16) = 17.49$, $p < .01$. Thirteen of the 16 children (81%) screened off and chose the domain-inappropriate cause (flipping the switch), whereas no child chose the domain-appropriate causes (the silly faces). By contrast, in the baseline condition, no child chose the domain-inappropriate causes, and 100% of the children chose the domain-appropriate causes.

Children's performance at ceiling in the baseline conditions emphasizes the fact that, consistent with past research, children's causal reasoning respects domain boundaries. Importantly, however, children's almost opposite performance in the test conditions indicates that a single trial providing evidence against domain-specific assumptions is sufficient for children to generate novel causal inferences.

Table 3
Number of Children per Type of Response in Experiment 4

Response	Physical effects		Psychological effects	
	Test	Baseline	Test	Baseline
Domain-inappropriate cause	12 (75)	0 (0)	13 (81)	0 (0)
Domain-appropriate causes	0 (0)	16 (100)	0 (0)	16 (100)
Other	4 (25)	0 (0)	3 (19)	0 (0)

Note. Numbers in parentheses are percentages. $n = 16$ per condition.

Significantly, the need to violate domain-specific assumptions had no effect on children's ability to make formal causal inferences. Given that the procedure for the test condition in Experiment 3 was formally identical to the one in Experiment 4, Experiment 3 could serve as a within-domain comparison. Children were as likely to screen off in the physical effects test condition (75% of children) and the psychological effects test condition (81% of children) of Experiment 4 as they were to screen off in the biology (77% of children) and psychology (67% of children) test conditions of Experiment 3.

Experiment 5

Experiment 4 suggested that children are able to use formal inference procedures, such as screening off, to override domain-specific knowledge. However, it is not clear from this experiment whether the children simply accepted that psychological causes could sometimes produce physical effects (and vice versa) or whether children used the evidence in the test condition to redefine the domain boundaries. For instance, the children might have determined (plausibly as it happens—although presumably 4-year-olds have limited experience with voice-activated technology) that spoken commands can be physical as well as psychological causes. Similarly, the children might have concluded that Catherine was an eccentric, easily amused by the activation of switches, rather than that Catherine's giggles were physically triggered by the switch.

The plausibility of such "expanded" domain specificity suggests both the difficulty of coming up with a principled way of defining a domain (see, e.g., Hirschfeld & Gelman, 1994) and the potential malleability of domain-specific knowledge. Children's knowledge might be domain specific, at least in part, because of the patterns of evidence children see rather than because of innate domain divisions. Specifically, children might generate domain-specific causal knowledge precisely to the extent that evidence for different causal principles in different domains is available in the input (i.e., to the extent that physical phenomena really do behave differently than biological or psychological phenomena). From this perspective, domain-specific causal inferences might be construed as defeasible outputs of, rather than constrained inputs to, more general learning mechanisms.

Children's willingness to extend their domain-inappropriate inferences might indicate whether domain-general inferences influence the development of domain-specific causal knowledge. If domain-specific knowledge acts as a strong constraint on children's causal inferences, children might be reluctant to generalize domain-inappropriate inferences to novel events. Instead, children might restrict their cross-domain inferences exclusively to the particular events for which they have evidence. Formal inference procedures might thus allow children to make exceptions to their domain-specific concepts without impacting their broader understanding of the domains.

Conversely, evidence and formal inference procedures might have an influence on the way that children construct domains of knowledge. If so, even minimal exposure to evidence that violates domain boundaries might significantly affect children's predictions about other events in the domains. Evidence might not just affect children's immediate causal judgments; it might transfer to causal predictions about novel events. Seeing that some causal relations can cross domain boundaries might make children more likely to predict that other relations can also cross those boundaries. Eventually this might lead children to redefine the domain boundaries themselves.

In order to assess how a conflict between domain-specific knowledge and formal inference procedures would influence children's new causal predictions, we modified Experiment 4 to permit a within-subject design. Children were first exposed to a screening-off test condition demonstrating cause–effect relationships comparable but not identical to those in Experiment 4. This allowed us to replicate the basic cross-domain result in Experiment 4 with new materials. Then children received a transfer task in which they were asked to make predictions about other events in the domains. This transfer task was identical to the baseline task in Experiment 4, but it took place after the children had been exposed to the screening-off test condition. Comparing the transfer task to the identical baseline task in Experiment 4 permitted us to assess whether children's predictions about novel causal events were influenced by exposure to evidence in which domain-appropriate causes were screened off by domain-inappropriate ones.

Method

Participants

Sixteen children ranging in age from 3 years 11 months to 5 years 4 months (mean age = 4 years 8 months) were recruited from an urban area preschool. Approximately equal numbers of boys and girls participated. Although most children were from White, middle-class backgrounds, a range of ethnicities reflecting the diversity of the population was represented.

Materials

Physical effects test task. A specially designed remote-operated light was used in the physical effects task. The light was in a 12 cm × 17 cm × 8 cm wooden box with an orange Lucite top. Two wires came out of the side of the box facing away from the child. One wire was noticeably plugged into a nearby wall socket. The other was attached to a concealed remote. When the concealed remote was put in the "on" position, the Lucite top would light up. When the remote was put in the "off" position, the light would turn off. None of the children gave any indication of suspecting the existence of the remote.

Two fake physical control devices, resembling switches, were also used. The fake controls (one white, one beige) had panels that flipped up and down to signal on and off. They were encased in 11 cm × 5 cm × 5 cm boxes. Both control boxes had electrical cords attached. The control devices looked functional but were not.

Psychological effects test task. Two pop-up puppets, a dog and a frog, were used in the psychological effects task. Each puppet was attached to a fuzzy green stick. When the stick was pushed up, the puppet popped out of a conical base. Both puppets had their tongues sticking out and a comical expression on their faces. Two cardboard paper towel tubes held upright in a cardboard box provided a stand for the puppets. A fake toggle was also used. This toggle was encased in a solid, white 11 cm × 4 cm × 5 cm box with a (nonfunctional) cord attached.

Figure 2 shows a drawing of the stimuli used in the test and transfer conditions of Experiment 5.

Transfer tasks. The same materials used in the warm-up task and the baseline condition of Experiment 4 were used in this condition.

Procedure

A female experimenter who was familiar to the children tested all the participants. In the test condition, a confederate assisted the experimenter. Children were brought into a quiet room in their school and sat facing the experimenter at a table. Children first participated in an unrelated experiment. In both the test and transfer conditions, order of domain presentation (physical effects vs. psychological effects) was counterbalanced across participants. Because the purpose of the study was to look at the impact of domain-inappropriate evidence on children's novel judgments, children always received the test condition first.

Test condition—physical effects task. For the physical effects task, the toy light was placed on the table and the children were told, "Some things make this toy light up. Can you help me figure out what makes the toy light up?"

This task was formally identical to the screening-off tasks in Experiment 4. The two domain-appropriate causes were the control devices; the domain-inappropriate cause was talking to the machine. The experimenter brought out each control device in turn (order of presentation was counterbalanced across subjects) and flipped the panel on and off three times. Nothing happened. The experimenter left each device in the "off" position and then looked at the toy and said, "Toy, turn on!" She surreptitiously activated the remote, and the toy lit up. While the toy was lit, the experimenter flipped Controls A and B back to the "on" position. After 10 s, the experimenter said to the child, "Can you make the light go out?"

Children saw that saying "Turn on!" and the machine lighting up were dependent even in the absence of Controls A and B, but that Controls A and B were independent of the effect in the absence of talking to the machine. If children are screening off, they should say something like "Toy, turn off!" Conversely, if the children are relying on domain-specific assumptions to guide their causal inferences and believe that physical causes (i.e., the control devices) are more likely than psychological causes (i.e., talking) to produce physical effects, then the children should flip the controls.

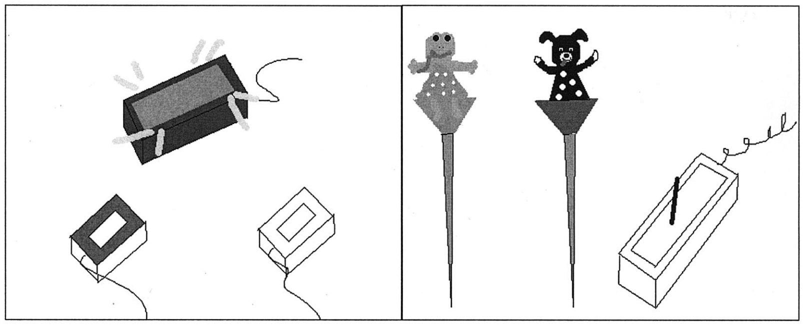
Test condition—psychological effects task. For the psychological effects task, the experimenter introduced the child to a confederate, saying, "This is my friend. My friend acts goofy sometimes. Can you help me figure out what makes my friend act goofy?"

This task was also formally identical to the screening-off tasks in Experiment 4. The two domain-appropriate causes were the frog and dog puppets; the domain-inappropriate cause was the toggle. The confederate closed her eyes, and the experimenter brought out each puppet in turn and placed it in the stand in view of the confederate (order of presentation was counterbalanced across subjects). Each time, the confederate did nothing. Then the experimenter brought out the toggle, counted "1, 2, 3," and flipped the toggle. The confederate immediately stuck out her tongue (like the puppets) and waggled her fingers in her ears. While she was "acting goofy," the experimenter returned Puppets A and B back to the stand within the confederate's line of gaze. After 10 s, the experimenter said, "My friend is acting goofy. Can you make her stop acting goofy?"

If children use screening-off assumptions to draw causal conclusions and are able to combine those assumptions with the substantive knowledge that undoing causes often removes effects, children should flip the toggle. Conversely, if the children rely on domain-specific assumptions and believe that psychological causes (i.e., goofy puppets) are more likely than physical causes (i.e., a toggle) to generate psychological effects, then the children should remove the puppets from the confederate's sight.

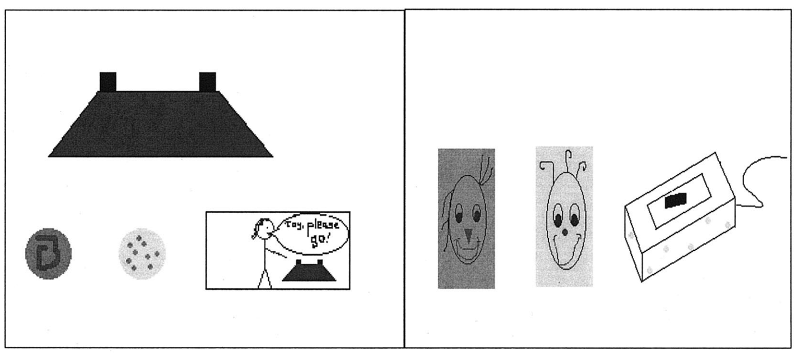
Transfer conditions. After the test condition, children participated in a warm-up task and a physical and psychological transfer condition identical to the baseline condition described in Experiment 4. As in that condition, children were asked to predict whether talking to the pottery wheel machine or placing a button on the machine, or both would cause it to make noise and whether silly faces or switches, or both would cause a person to giggle.

If domain-specific knowledge acts as a relatively strong constraint on children's causal inferences, then children in the transfer condition in



Physical effects test condition: toy light, and the two controls.

Psychological effects test condition: the "goofy" puppets and the toggle.



Physical effects transfer condition: noise-making machine, the two buttons, and the picture of talking to the machine.

Psychological effects transfer condition: the two silly faces and the switch.

Figure 2. Materials used in Experiment 5.

Experiment 5 should perform identically to the children in the baseline condition in Experiment 4. If, however, formal inference mechanisms such as screening off influence children's understanding of domains, then the children in Experiment 5 might generalize from the evidence in the test condition, and their transfer responses might be significantly different from those of the children in the baseline condition in Experiment 4.

Results

Preliminary analyses revealed no effect of order of domain presentation on the children's responses in either the test or baseline conditions. In both the test and transfer conditions, the children could choose the domain-inappropriate cause, one or both domain-appropriate causes, or something else altogether (i.e., they could make no response or say "I don't know"). Children's responses are presented in Table 4.

Physical Effects Tasks

Within the physical effects test condition, children were significantly more likely to choose the domain-inappropriate cause than to choose the domain-appropriate causes or to make another response, $\chi^2(2, N = 16) = 3.77, p < .05$. Ten of the 16 children (62%) screened off and chose the domain-inappropriate cause (i.e., they said, "Turn off" or "Turn off now"), whereas only 3 children (19%) chose the domain-appropriate causes (the controls). By contrast, in the transfer condition, children were significantly more likely to choose the domain-appropriate causes than to choose the domain-inappropriate cause or to make another response, $\chi^2(2, N = 16) = 14.09, p < .01$. Twelve of the 16 children (75%) chose the domain-appropriate causes (the buttons), whereas 4 children (25%) chose the domain-inappropriate cause (the picture of the experimenter talking to the machine).

Critically, however, children were significantly more likely to choose the domain-inappropriate causes in the transfer condition

Table 4
Number of Children per Type of Response in Experiment 5

Response	Physical effects		Psychological effects	
	Test	Baseline	Test	Baseline
Domain-inappropriate cause	10 (62)	4 (25)	13 (81)	5 (31)
Domain-appropriate causes	3 (19)	12 (75)	0 (0)	11 (69)
Other	3 (19)	0 (0)	3 (19)	0 (0)

Note. Number in parentheses are percentages. $N = 16$.

of Experiment 5 than in the baseline condition of Experiment 4, $\chi^2(1, N = 32) = 4.57, p < .05$. Twenty-five percent of the children chose domain-inappropriate causes in the transfer condition of Experiment 5 (3 children chose only the inappropriate causes, and 1 child chose all three causes), whereas no children chose domain-inappropriate causes in the baseline condition of Experiment 4.

Psychological Effects Tasks

Within the psychological effects test condition, children were also significantly more likely to choose the domain-inappropriate cause than to choose the domain-appropriate causes or to make another response, $\chi^2(2, N = 16) = 14.09, p < .01$. Thirteen of the 16 children (81%) screened off and chose the domain-inappropriate cause (i.e., they flipped the toggle), whereas no children chose the domain-appropriate causes (the puppets). By contrast, in the transfer condition, children were significantly more likely to choose the domain-appropriate causes than to choose the domain-inappropriate cause or to make another response, $\chi^2(2, N = 16) = 11.45\ p < .01$. Eleven of the 16 children (69%) chose the domain-appropriate causes (the silly faces), and 5 children (31%) chose the domain-inappropriate cause (the switch).

Again, however, children were significantly more likely to choose the domain-inappropriate causes in the transfer condition of Experiment 5 than in the baseline condition of Experiment 4, $\chi^2(1, N = 32) = 5.92, p < .03$. Thirty-one percent of the children chose domain-inappropriate causes in the transfer condition of Experiment 5 (4 children chose only the domain-inappropriate cause, and 1 child chose all three causes), whereas no children chose domain-inappropriate causes in the baseline condition of Experiment 4.

Discussion

Overall, the findings of Experiment 5 replicate the findings of Experiment 4, suggesting that children are aware of domain boundaries but can use formal principles, such as screening off, to override them. The difference between the transfer condition in Experiment 5 and the identical baseline condition in Experiment 4 suggests that children's causal predictions about new events are influenced by exposure to evidence in which domain-inappropriate causes screen off domain-appropriate ones. Importantly, however, most children continued to respect domain boundaries in the transfer condition despite the similarity of the stimuli in the test and transfer conditions and despite children's willingness to override domain boundaries in the test condition.

Thus, the results of Experiment 5 suggest that domain-specific knowledge and domain-general formal principles interact in influencing children's causal judgments. When children do not have any specific evidence about a set of events, they make judgments that are based on domain-specific prior knowledge. When children are given evidence about those events, the significant majority of children make judgments that are based on the evidence, even when that evidence contradicts domain-specific knowledge. However, when children are given domain-inappropriate evidence about one set of events and are asked to make predictions about novel though similar events, they behave in an intermediate way. In this case, children do not immediately overthrow domain-specific knowledge in the face of novel evidence; the majority of children make judgments that are consistent with their prior knowledge. Nonetheless, in Experiment 5, a sufficient number of children were influenced by the counterintuitive evidence to change the overall pattern of responses. In combination with prior knowledge, formal inference procedures thus might allow for learning that is both conservative and innovative.

It seems probable that the similarity between the test and transfer causes in Experiment 5 influenced children's willingness to generalize from conditions for which they had evidence to those for which they did not. The target causes (saying "turn on" and "please go"; switches and toggles) and the effects (light and noise; goofiness and giggling) were designed to be quite comparable. Children might have been less willing to make the transfer had the stimuli been less similar. On the other hand, children had very little exposure to the cross-domain effects. They were given only a single trial with a single set of materials. Children might have been more willing to make the transfer with more exposure to evidence. Thus, although this study suggests that formal inferential mechanisms can influence children's understanding of domains, additional research must determine the degree of that influence.

General Discussion

The results of these five studies suggest that young children can make causal judgments using patterns of independent and dependent probabilities across a range of tasks and domains. Earlier research demonstrated that 3- and 4-year-old children could use the principles of screening off to identify causes and craft new interventions in the domain of physical causality (Gopnik et al., 2001). The present research suggests that preschool children are also able to use screening-off information to learn the causal structure of biological and psychological events. Moreover, children's ability to draw causal inferences from patterns of dependence is quite robust. The more demanding procedures used in Experiments 2 and 3 and the introduction of domain-inappropriate causes in Experiments 4 and 5 in no way impaired children's ability to infer causal structure from patterns of dependence.

One interesting feature of this study is the lack of domain differences. Across all the domains, children distinguished between screening-off trials and trials in which only simple frequency information was available, and children also distinguished between screening-off trials and trials in which there were multiple independent causes. Furthermore, children used screening-off information both in tasks requiring them merely to identify the cause of events and in tasks requiring them to develop novel interventions. The consistency of results across tasks and domains suggests that the particular content of any given event is less important to causal inference than are the independent and dependent probabilities of the events.

Children's inferences about causes that cross the boundaries of domains are particularly interesting. Although these studies replicate previous research in demonstrating that children reason in a domain-appropriate manner, children performed as well on cross-domain screening-off tasks as on tasks within a single domain. Children's considerable domain-specific knowledge did not appear to constrain their ability to make formal causal inferences in violation of domain-appropriate ones. Indeed, the converse appeared to be true: Children's formal causal inferences influenced their baseline judgments about domain-specific causality.

However, although these findings suggest that formal causal learning mechanisms are quite general, and might even contribute to the development of domain-specific theories, domain-specific knowledge might nonetheless influence children's causal judgments. Specifically, the transfer findings in Experiment 5 suggest that prior domain-specific knowledge might limit children's willingness to generalize inferences that are inconsistent with such knowledge. A tension of this sort, between children's prior beliefs and their assimilation of novel information, is consistent with a theory theory approach to conceptual development (see, e.g., Gopnik & Meltzoff, 1997). This sort of interaction between prior causal knowledge and new causal inferences can also be captured in the Bayes net formalism. In combination with domain-specific knowledge, formal inference procedures might provide a mechanism for learning that would allow children's causal theories to be both stable and defeasible.

In addition, it seems probable that innate or very early developing domain-specific concepts might provide a necessary foundation for formal causal reasoning. It is hard to imagine, for instance, how one might track the dependence and independence of objects and events without some initial knowledge of what constitutes an object. Similarly, it seems improbable that children could develop inferences about the causes of human behavior without domain-specific abilities to process facial expressions, to recognize emotion, and so forth. It is even possible that children are born with some causal assumptions. However, these assumptions may be overridden later if they are not congruent with the evidence. The exact nature of the dynamic between domain-general causal learning mechanisms and domain-specific concepts remains an area for further research.

Note also that although formal causal learning mechanisms such as screening off may be quite general in that they can be applied to information from many types of domains, such formal mechanisms are nonetheless more constrained than traditional domain-general learning mechanisms, such as logical inference or associations. Causal learning may be quite general, in the sense that it can act on information from a variety of domains and infer even novel mechanisms as possible causes, while still being constrained by assumptions about how evidence and causal structure are related. This might include the assumptions that events have causes, that causes produce effects, that effects do not produce causes, and that intervening on a cause will influence the effect, or it might include the more general formal assumptions embodied in the causal Markov and faithfulness assumptions. These assumptions, unlike logical assumptions, are contingently rather than necessarily true and, unlike associationism, are specific to domains where causal structure can be recovered. The process of causal learning is thus quite different from either the process of deductive inference or the process of capturing high-level regularities in the input, as in classic associationist accounts.

Finally, it should be stressed that there are multiple models by which patterns of independent and dependent probabilities could lead to causal inferences. We have discussed the relationship between causality and probabilistic dependence given in the literature on causal Bayes nets (Pearl, 1988, 2000; Spirtes et al., 1993). According to the formalism, only some causal graphs, and not others, are compatible with particular patterns of conditional probability among events, given the causal Markov and faithfulness assumptions. The children in our studies did in fact infer the correct graphs given the data. However, even within the Bayes net literature, there are many different specific algorithms that might be used to derive the correct causal graph from screening-off data. In some of these algorithms, the graphs are derived directly from the patterns of independent and dependent probabilities. Other algorithms rely on Bayesian methods of assigning prior probabilities to all the possible graphs and computing their posterior probabilities given the data (see, e.g., Glymour & Cooper, 1999). Cheng's "causal power" model is also a special case of a causal Bayes net model, and her learning rules might be construed as Bayes net learning algorithms (Cheng, 1997, 2000). It might also be possible to use the Rescorla–Wagner learning rule (Rescorla & Wagner, 1972) as the basis for a restricted learning algorithm in a causal Bayes net account, though such an account would require assumptions that go well beyond the Rescorla–Wagner rule itself. The present experiments do not test which of these specific accounts best explains the causal judgments of young children; however, see Gopnik et al. (2004) for further discussion and data discriminating among these accounts.

What we have shown is that across domains, even very young children can and do infer new causal relations from information about dependent and independent probabilities. Moreover, they are able to integrate these formal inductive inferences with substantive and domain-specific knowledge. Work in computer science and artificial intelligence has shown that computational learning mechanisms can produce normative causal inferences. Work in developmental psychology has demonstrated that young children are able to learn the causal structure of events with remarkable speed and accuracy. Future investigations may help us to integrate these fields of research and to better understand the mechanisms that make causal learning possible.

References

Bartsch, K., & Wellman, H. (1995). *Children talk about the mind.* New York: Oxford University Press.

Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review, 77,* 32–48.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York: Academic Press.

Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169–200). New York: Cambridge University Press.

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

Cheng, P. W. (2000). Causality in the mind: Estimating contextual and conjunctive power. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 227–253). Cambridge, MA: MIT Press.

Chomsky, N. (1981). Knowledge of language: Its elements and origins. *Philosophical Transactions of the Royal Society of London, 295*(1077, Series B), 223–234.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1995). Young children's knowledge about thinking. *Monographs of the Society for Research in Child Development, 60*(1, Serial No. 243), v–96.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essence: Early understandings of the non-obvious. *Cognition, 38,* 213–244.

Glymour, C. (2001). *The mind's arrows: Bayes nets and causal graphical models in psychology.* Cambridge, MA: MIT Press.

Glymour, C., & Cooper, G. F. (1999). *Computation, causation, and discovery.* Cambridge, MA: MIT Press/AAAI Press.

Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111,* 1–30.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts and theories.* Cambridge, MA: MIT 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.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293). New York: Cambridge University Press.

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

Hickling, A. K., & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology 37,* 668–683.

Hirschfeld, L. A., & Gelman, S. A. (1994). Toward a topography of mind: An introduction to domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 3–35). New York: Cambridge University Press.

Inagaki, K., & Hatano, G. (1993). Young children's understanding of the mind–body distinction. *Child Development, 64,* 1534–1549.

Kalish, C. (1996). Causes and symptoms in preschoolers' conceptions of illness. *Child Development, 67,* 1647–1670.

Keil, F. C. (1995). The growth of causal understandings of natural kinds. In D. Sperber & D. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (Symposia of the Fyssen Foundation, pp. 234–267). New York: Clarendon Press/Oxford University Press.

Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). New York: Cambridge University Press.

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

Notaro, P. C., Gelman, S. A., & Zimmerman, M. A. (2002). Biases in reasoning about the consequences of psychogenic bodily reactions: Domain boundaries in cognitive development. *Merrill-Palmer Quarterly, 48,* 427–449.

Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. *Cognitive Development, 5,* 193–207.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems.* San Mateo, CA: Morgan Kaufmann.

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

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Reichenbach, H. (1956). *The direction of time.* Berkeley: University of California Press.

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: Appleton-Century-Crofts.

Schwartz, B. (1978). *Psychology of learning and behavior.* New York: W. W. Norton.

Seligman, M. E., & Johnston, J. C. (1973). A cognitive theory of avoidance learning. In F. J. McGuigan & D. B. Lumsden (Eds.), *Contemporary approaches to conditioning and learning* (pp. 69–110). Oxford, England: V. H. Winston.

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

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 21, pp. 229–261). San Diego, CA: Academic Press.

Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development, 47*(1), 1–51.

Sidman, M. (1953). Avoidance conditioning with brief shock and no exteroceptive warning signal. *Science, 118,* 157–158.

Sobel, D. M. (2001, April). *Towards a "theories" account of counterfactual reasoning.* Paper presented at the 2001 biennial meeting of the Society for Research in Child Development, Minneapolis, MN.

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

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99,* 605–632.

Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7,* 337–342.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search* (Springer Lecture Notes in Statistics). New York: Springer-Verlag.

Tenenbaum, J. B., & Griffiths, T. L. (in press). Theory-based causal inference. In S. Thrun, L. Saul, & B. Schölkopf (Eds.), *Advances in neural information processing systems 16* (Proceedings of the 2003 Neural Information Processing Systems Conference). Cambridge, MA: MIT Press.

Wellman, H. M., Hickling, A. K., & Schult, C. A. (1997). Young children's psychological, physical, and biological explanations. In H. M. Wellman & K. Inagaki (Eds.), *The emergence of core domains of thought: Children's reasoning about physical, psychological, and biological phenomena* (New Directions for Child Development No. 75, pp. 7–25). San Francisco: Jossey-Bass/Pfeiffer.

Woodward, J. (2002). *Making things happen: A theory of causal explanation.* New York: Oxford University Press.

Received August 5, 2002
Revision received October 2, 2003
Accepted October 8, 2003 ■