L. Schulz, A. Gopnik, & C. Glymour (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science (special section on Bayesian and Bayes-Net approaches to development)* 10, 3, May 2007, pp. 322-332.

Preschool children learn about causal structure from conditional interventions

Laura E. Schulz[1], Alison Gopnik[2] and Clark Glymour[3]

[1] Massachusetts Institute of Technology, [2] University of California, Berkeley, [3] Carnegie Mellon University and Institute for Human and Machine Cognition, Pensacola, Fl

This research was supported by an American Association of University Women Dissertation Fellowship and a National Science Foundation Award (#99-136) to L.S, a National Science Foundation grant (DLS0132487) to A.G, and grant NN00014-03-1-0516 from the Office of Naval Research to IHMC to C. G, and a James S.McDonnell Foundation Causal Learning Collaborative Initiative grant to all three authors. We are grateful to Tamar Kushnir, Christine Schaeffer, and Wendy Wienerman for helpful comments and suggestions. Thanks also to Fruit and Flower, Northeast Community Child Development Center, Helen Gordon, the Goddard School, Willowbrook, and the Berkeley Child Study Center. Address for correspondence: L. Schulz, NE20-459, MIT Department of Brain and Cognitive Sciences, 77 Massachusetts Avenue, Cambridge, MA 02139. Phone: 617-253-7957. E-mail: lschulz@mit.edu

Developmental Science. Vol. 10(3), May 2007, pp. 322-332

Preschool children learn causal structure from conditional interventions

Abstract

The conditional intervention principle is a formal principle that relates patterns of interventions and outcomes to causal structure. It is a central assumption of experimental design and the causal Bayes net formalism. Two studies suggest that preschoolers can use the conditional intervention principle to distinguish causal chains, common cause and interactive causal structures even in the absence of differential spatiotemporal cues and specific mechanism knowledge. Children were also able to use knowledge of causal structure to predict the patterns of evidence that would result from interventions. A third study suggests that children's spontaneous play can generate evidence that would support such accurate causal learning.

Preschool children learn causal structure from conditional interventions

Until recently, research on children's causal learning has focused primarily on children's understanding of causal mechanisms. Even infants use spatiotemporal cues to infer contact causality (e.g., Leslie & Keeble, 1987) and young children understand many domain-specific causal relations (Ahn, Gelman, Amsterlaw, Hohenstein, & Kalish, 2000; Bullock, Gelman, & Baillargeon, 1982; Carey & Spelke, 1994; Shultz, 1982; Spelke, Breinlinger, Macomber, & Jacobson, 1992). In adult cognitive psychology by contrast, researchers have focused primarily on domain-general causal learning from associations (Shanks, 1985; Shanks & Dickinson, 1987; Spellman, 1996) and patterns of covariation (Cheng, 1997, 2000) among events.

Recently, however, researchers have suggested that the crucial piece missing from both mechanism and covariation accounts of causal reasoning is the notion of intervention. Specifically, researchers have suggested that knowing that X directly causes Y means knowing that, all else being equal, intervening to change X can change Y (Pearl, 1998; 2000; Spirtes, Glymour & Scheines. 1993; Woodward, 2003). We will discuss the interventionist account of causation first intuitively and then formally with respect to its role in causal Bayes net learning algorithms.

Suppose you notice that when you go to a party and drink wine, you do not sleep well. This could be because wine is keeping you up, and parties make you drink wine, or because parties both cause insomnia and make you drink wine. Assuming that these are the only three relevant variables, you could learn the correct causal structure by intervening to hold one factor constant and then intervening to vary the other factor. For example, you could try going to parties and drinking and going to parties sober. If there is no difference in how you sleep, then wine is not likely to be the direct cause of your insomnia. If there is a difference, then wine is a likely cause -- and you can make this inference even if you do not know anything about the mechanism by which wine keeps you awake.

Causal relations, like the relationship between drinking, partying and insomnia, can be represented by directed graphs, also called causal Bayes nets (see Figure 1, and introduction to this special section for more details). Events (e.g., wine, partying, insomnia, etc) are represented by variables that can take particular values (e.g., present or absent). Relations between the variables are represented by directed edges (arrows) connecting those variables. The structure of the graph constrains the conditional probabilities of the variables. All variables in a graph are assumed to be probabilistically independent of all other variables except their own descendents, conditional on their own immediate ancestors.

Insert Figure 1 here

Critically, in a *causal* Bayes net, the arrows encode not just constraints on probabilistic dependencies but causal relationships. In particular, arrows between variables imply that interventions on the variable on the left side of the arrow will lead to changes in the variable on the right side. Within the formalism, interventions are treated as special additional variables with special features: (i) they are exogenous (that is they are not influenced by any other causal factors in the graph); (ii) they fix the probability distribution of the variables of interest; and (iii) they influence other variables in the graph only through their effect on the intervened on variable. If you know the structure of the causal graph, you can infer the outcome of interventions. Conversely, if you do not know the graph, you can use data from interventions to learn the underlying structure.[1]

One way to capture these relations between interventions, dependencies and causal arrows formally is as follows: for a set of variables in a causal graph, W directly causes I (W→I) if and only if: (i) An intervention could fix the values of all other variables in the graph and result in I having a particular probability distribution (P(I)) such that (ii) another intervention on the value of W (iii) will change the probability distribution of I from P(I) to P'(I) but (iv) not influence I other than through W and (v) not change the fixed value of the other variables in the graph.[2] We will call this the conditional intervention principle.

Although this principle might sound complex, it is simply a formal statement of the intuitions that underlie experimental design. To find the causal relationship between two variables, you can intervene to fix the distribution of potential confounders and then see what happens when you intervene to manipulate the variable of interest. If, controlling for all else, changing the value of W changes the value of I, we can conclude that there is a direct causal relationship between W and I – and we can learn this relationship even if we do not know the underlying mechanism. Conversely if, controlling for all else, changing the value of W fails to change the value of I, we can

[1] Throughout, we use the term *inference* to mean predicting the data generated by a known causal structure and *learning* to mean using observed data to discover the underlying causal structure.
[2] Although we deal with actual human interventions in this paper, "natural experiments" and counterfactual interventions are also possible (see Woodward, 2003).

conclude that the relationship is not causal, even if the variables are correlated. Unlike associationist and covariation accounts of causal learning (Cheng, 1997; 2000; Shanks & Dickinson, 1997), Causal Bayes nets distinguish evidence obtained by observation from and evidence obtained by intervention and make different inferences from the two types of evidence.

In this study, we look at whether children can use the conditional intervention principle to learn the causal structure underlying observed data. Imagine for instance, that you flip a switch and two gears start to spin simultaneously. Assuming the causal relationships between the gears are generative (i.e., rather than inhibitory) and deterministic, there are four possibilities: a) A makes B spin b) B makes A spin c) Neither gear makes the other spin (because the switch independently makes each gear spin) d) Neither gear will spin without the other (because the switch and A together make B spin and the switch and B together make A spin). The pictures in Figure 2 represent these alternatives.

Insert Figure 2 here

If you could remove gear A and see if gear B stops (or vice versa), then you could learn the causal relation between A and B from the immediate effect of your own interventions. But suppose (perhaps because your fingers might get pinched) you cannot remove a gear when the switch is on. You could still distinguish the structures by observing the conditional dependence among interventions and outcomes. You could for instance, remove gear A, flip the switch on, and observe gear B. Then you could flip the switch off, replace gear A, remove gear B and flip the switch on (see Figure 3). Note that these interventions do not change the association between the two gears. Also note that the effect of the intervention to remove A is not to stop B (because B isn't moving in the first place). Nor is the effect of the intervention to replace A, to make B spin (because B is still in both cases). However, consistent with the conditional intervention principle, for a fixed value of other causes of B (the switch), an intervention on A changes the value of B (compare for instance lines 5 and 6 of the table in Figure 1a) whereas for no value of the switch will an intervention on B change the value of A (compare lines 3 and 4 and lines 7 and 8). You should conclude that A→B.

Insert Figure 3 here

Recall that in a causal Bayes net, arrows between variables specify how interventions to change the value of one variable will affect the probability distributions of other variables. Because the gear system is deterministic, the causal Bayes net probability predictions reduce to simple deterministic equations; the value of each variable, X can be expressed as a function of the values of the variables whose arrows point into X. Each equation corresponds to a particular causal structure (shown in Table 1). Thus you can learn the causal structure of the system by observing the data, determining which equation could generate that data, and mapping that equation on to the corresponding structure.

In the table, each variable has two possible values. S = 1 means S is on and S = 0 means S is off; A = 1 means A spins and A = 0 means A does not spin (similarly for B).

Intervening on a variable forces it to have a particular value. For example, removing a wheel from its spindle forces the value of the wheel to be “not spinning”, (Pearl, 2000). Thus if the structure is S→A→B, and we intervene on S, then S will take the value it’s set to, A will take the value of S, and B will take the value of A (i.e., row 1, column 2). If A is removed and forced not to spin (that is, A = 0), the other equations still hold; B will take the same value as A (column 3). However, if B is set to 0, A will be unaffected because A = S (column 4). On the other hand, if the structure is S→B-→A a different pattern of data will result. Thus the computational formulas for the causal graphs produce the relation between structure and evidence captured by the conditional intervention principle. Because each structure corresponds to a different formula, which generates a different pattern or evidence, the structure can be uniquely determined by the data.

This type of inference is not easily explained by other accounts of causal learning. There is a plausible mechanism underlying each of these relationships and the gears’ movements are simultaneous and spatially contiguous in all cases. Many researchers have suggested that we evaluate the strength of candidate causes based on their covariation with a known effect (Cheng, 1997; 2000; Novick & Cheng, 2004; Shanks & Dickinson, 1987). However, in this case the gears could be either causes or effects. Indeed, causal learning very often requires determining whether X causes Y or Y causes X. Covariation models do not explain such learning.

Recently causal Bayes nets have been used to model a variety of causal reasoning problems (Gopnik & Schulz in press) both in adults (Glymour, 2001; Lagnado & Sloman, 2002; Rehder & Hastie, 2001; Steyvers, Tenenbaum, Wagenmakers & Blum, 2003; Tenenbaum & Griffiths, 2001; Waldmann & Hagmayer, 2001) and children (Gopnik, Sobel, Schulz, & Glymour, 2001; Gopnik et al., 2004; Schulz & Gopnik, 2004, ; Sobel, Tenenbaum & Gopnik, 2005; Kushnir & Gopnik, 2005; in press). Some research suggests that children can use information about interventions to make inferences about the direction of simple causal relations (Gopnik, et al., 2004). However, as yet there have been no studies looking at whether children or adults can use evidence consistent with the conditional intervention principle to learn more complex causal structures.

Moreover, even if children can make accurate inferences given informative evidence, it is not clear how children might get such evidence (outside the lab). Considerable research suggests that both children and adults are poor at designing informative experiments (Chen & Klahr, 1999; Inhelder & Piaget, 1958; Kuhn, 1989; Kuhn, Amsel & O'Laughlin, 1988; Masnick & Klahr, 2003). However, it is possible that in the course of free play, children might spontaneously generate evidence that could support causal learning; children might then implicitly perform accurate computations given the patterns of evidence. In the following three experiments, we investigate children's ability to learn causal structures from the outcome of interventions, children's ability to infer the outcome of interventions from knowledge of causal structure, and the extent to which children's spontaneous play generates evidence that could support accurate causal learning.

Experiment 1

Procedure

Participants.

Seventy preschool children (mean age: 56 months; range: 42– 66 months) were assigned to one of four conditions: Test Condition 1 (n = 20); Test Condition 2 (n = 20);

Test Condition 3 (n = 14), and a Switchless Control condition (n = 16). In this study and the following two studies, approximately equal numbers of boys and girls participated (55% girls). The majority of children were white and middle class but a range of ethnicities consistent with the diversity of the local population was represented.

Test conditions.

A custom-built electronic toy was used (see Figure 4). The toy had two pegs; each could support one of four uniquely colored gears. A new pair of gears was used on every trial. Sensors detected the presence of the gears and a hidden control allowed the experimenter to implement each of the four structures in Figure 1. A switch on the front activated the toy. Two sets of four pictures like those in Figure 1 were also used.

Insert Figure 4 here

Trials began with the switch in the off position. The experimenter placed and removed each gear on the toy in turn, explaining that she could "take the gears on and off the machine". She flipped the switch on to make both gears spin simultaneously and flipped the switch off so that both gears were still. This provided evidence equivalent to lines 1 – 4, 5, and 7 in Figure 1.

Pilot work suggested that children could not keep track of four pictures simultaneously so the children were tested in three conditions in order to exhaustively compare their ability to distinguish the structures. In each condition, three pictures were set before the child in random order. Children were told, "Here are some ways the toy could work". Children in Test Condition 1 saw pictures corresponding to 1a, 1b, and 1d.

Children in Test Condition 2 saw pictures corresponding to 1a, 1b, and 1c. Children in Test Condition 3 saw pictures corresponding to 1a or 1b (particular picture counterbalanced between subjects), 1c, and 1d. The pictures were colored to match the gears on the toy on each trial and were described in terms of the colors of the gears. For example, for yellow and green gears, and 1a, children were told, “This picture shows that yellow is pushing green. Yellow makes green go” (and the reverse for 1b); for 2c they were told, “Green doesn’t push yellow and yellow doesn’t push green. They each push themselves.” For 1d they were told: “Green pushes yellow and yellow pushes green. Both wheels push together.” Children were asked to redescribe each picture and corrected if necessary.

The children received two trials in counterbalanced order. In all conditions, the experimenter removed and replaced each gear in turn, showing the children how the toy behaved when both gears were present and when each gear was isolated. In Test Condition 1, one trial provided evidence for structure 1a, the other for 1d. In Test Condition 2, one trial provided evidence for structure 1b, the other for 1c. In Test Condition 3, one trial provided evidence for structure 1c, the other for structure 1d. For instance, for structure 1a the experimenter removed gear A, turned on the switch, and B failed to spin. She turned off the switch, replaced A, flipped the switch and both gears spun. She then removed gear B, turned on the switch, and A spun (see Figure 3). This procedure provided the information in lines 6 and 8 of Figure 1a. A new pair of gears was used on each trial, and the toy was rotated 90° between trials to avoid any position biases. After each trial, children were asked, “Can you give me the picture that shows how the toy is working right now?”

Switchless Control Condition.

Children in the Control Condition received the same evidence as children in Group 1. However, before any gears were removed, the toy was turned 180° so children couldn't see the position of the switch. This manipulation preserved the perceptual and associative relation between the gears. If children can use these cues to identify the correct causal structure, they should respond like children in Group 1. Formally, however, the evidence is confounded: a gear might not spin either because a causal gear was removed or because the switch was off. If children are relying on the pattern of interventions, they should not uniquely prefer a causal structure in this condition.

Results and Discussion

Preliminary analyses revealed no effect of trial order. In all three Test Conditions and on both trials, children were significantly more likely to choose the correct picture than expected by chance and to choose the correct picture more than any other picture. The results from the four conditions are presented in Table 2.

Insert Table 2 here

In Condition 1, given evidence for 1a, children chose picture 1a above chance (n = 20, $p < .001$ by binomial test) and more than 1b (n = 16, $p < .001$ by binomial test) or 1d (n = 19, $p < .025$ by binomial test). Given evidence for 1d, children chose picture 1d above chance (n = 20, $p < .001$ by binomial test) and more than 1a (n=19, $p < .001$ by binomial test) or 1b (n = 18, $p < .001$ by binomial test). In Condition 2, given evidence for 1b, children chose picture 1b above chance (n = 20, $p < .005$ by binomial test) and more than 1a (n = 16, $p < .025$ by binomial test) or 1c (n = 17, $p < .025$ by binomial test). Given evidence for 1d, children chose picture 1d above chance (n = 20, $p < .001$ by binomial test) and more than 1a (n=19, $p < .001$ by binomial test) or 1b (n = 18, $p < .001$ by binomial test). Children chose correctly on both trials significantly more often than expected by chance: 60% in Condition 1 (n = 20, $p < .001$ by binomial test) and 55% in Condition 2 (n = 20, $p < .001$ by binomial test).

Similarly, in Condition 3, given evidence for 1c, children chose picture 1c above chance (n = 14, $p < .005$ by binomial test) and more than 1d or the chain (n = 12, $p < .05$ by binomial test for both). Given evidence for 1d, children chose picture 1d above chance (n = 14, $p < .001$ by binomial test) and more than 1c or the chain (n = 14, for 1c $p < .025$ by binomial test); no children chose the chain. Fifty-seven percent of the children chose the correct structure on both trials, significantly above chance (n = 14, $p < .001$ by binomial test).

These results suggest that children can use evidence consistent with the conditional intervention principle to distinguish chains from each other and from common effects and conjunction structures. They did so even in the absence of distinguishing mechanistic cues; the gears moved simultaneously and interlocked so there were no perceptual cues to the causal structure. Note that if children had used simpler heuristics (e.g., spinning wheels push other wheels; still wheels do not push other wheels) they would not have made correct judgments about structures 1c or 1d. Note also, that even though the common effects and conjunction structures conflict with adult knowledge about gear mechanisms, preschoolers did not seem to find these structures more difficult.

It is possible that children might have assumed that no gear could move spontaneously. Thus if the switch failed to move a gear on its own, the children might have inferred that the other gear must push it. However, if children used this heuristic to identify the causal structure, they could have chosen the interactive structure for all conditions (i.e., this assumption might explain why children believe that gears that do not move with the switch *are* pushed but they do not explain why children believe that gears that do move with the switch *are not* pushed).[3]

In the Switchless Control Condition, children also chose among the structures at chance. Given evidence about the gears (but not the switch) comparable to 1a, children chose picture 1a at chance and likewise for 1d (n = 16, $p = ns$ by binomial test for each). These results suggest that the relative salience of the gears or the presence of a single moving gear were not sufficient cues for children to make accurate causal inferences. Overall, these results are consistent with the possibility that children can use evidence from interventions to learn a wide range of causal structures

Experiment 2

A central feature of causal Bayes net learning algorithms is that they work in both directions. You can use evidence from interventions to learn causal structure but you can also use knowledge of causal structure to infer the patterns of evidence that will result

[3] Additionally in a related study (Schulz, 2003) children were given partial evidence about the causal structures (i.e., evidence about one gear but not the other). Eighteen children were shown that both gears spun together and then shown that the switch made one of the gears spin by itself; 18 were shown the gears spinning together and then shown that the switch failed to make a single gear spin. Even given a forced choice between gears, children were unable to use the partial evidence to identify the causal gear. This suggests that children rely on the full set of information about interventions and outcomes to disambiguate the possible causal structures rather than simply assuming that still gears are pushed.

from interventions. In Experiment 2, we told the children the causal structure and looked at what inferences they could make about the outcome of interventions.

Procedure

Participants.

Sixteen children (mean age: 54 months; range: 42– 61 months) were tested.

Training.

The experimenter introduced the gear toy to the child. The experimenter placed and removed each gear on the toy in turn, explaining that she could "take the gears on and off the machine". She placed one gear on a peg (left/right peg counterbalanced between children) flipped the switch on, and the gear spun. The experimenter explained, "Some gears spin by themselves." She removed that gear, surreptitiously reset the toy, and put a new gear on the same peg. She flipped the switch on and the gear stayed still. The experimenter explained, "Some gears do not spin by themselves." The training familiarized children with the toy's function and affordances.

Test.

The experimenter brought out two new gears and set them on the toy. She flipped the switch on so that both wheels spun and then flipped the switch off. The experimenter held up a picture and said, "This shows what is happening on the toy right now." The pictures were described as in Experiment 1. Children received four trials, one for each of the four pictures in Figure 1 (order counterbalanced). On each trial, the experimenter placed a picture in front of the child, removed the left gear and held the right gear above its peg. She said, "If I put this gear down right now and turn on the switch, will the gear spin or the gear stay still?" The experimenter repeated this with the right gear.

Results and Discussion

Preliminary analyses revealed no order effects. Children were counted as answering correctly only if they answered appropriately for both gears (i.e., for 1a, still/spin; for 1b, spin/still; for 1c, spin/spin; for 1d, still/still). Children's responses are shown in Table 3.

Insert Table 3 here

For each of pictures, 1a, 1b, and 1d, children made the correct response significantly above chance (n = 16, $p < .05$ by binomial test for each); for picture 1c, there was a trend for children to respond correctly (n = 16, $p = .08$ by binomial test). Other than the correct response, no pattern of responding approached significance for any structure. Twelve percent of the children made the correct predictions on all four trials, significantly more than expected by chance (n = 16, $p < .001$ by binomial test).

The results of Experiment 4 suggest that children can use knowledge of causal structure to predict the pattern of evidence that will result from interventions. Knowing the causal relationship between the gears, children were able to predict how an intervention on the switch would affect each gear individually. Consistent with the causal Bayes nets formalism, these results suggest that children can use knowledge of causal structure to predict the patterns of conditional dependence and independence that result from interventions.

Experiment 3

In the preceding experiments, we controlled the information children received. That is, we isolated the variables and showed the children how each gear behaved individually. Would children be able to generate this evidence by themselves and accurately learn the causal structure from the evidence of their own interventions? We did not expect that children would design an appropriate experiment to learn the causal structure (i.e., that they would deliberately remove each gear in turn and test the remaining gear with the switch); however we believed preschoolers might, in the course of play, spontaneously produce the type of evidence that would support accurate causal learning. We tested children both singly and in dyads because we believed that children playing in pairs might generate a broader range of interventions (and thus be more likely to generate informative evidence) than children playing by themselves.

Procedure

Participants.

Forty children (mean age: 57 months; range: 49– 66 months) were tested. Twelve children were assigned to a Single Causal Chain condition; 12 to a Dyad Causal Chain condition, and 16 to a Dyad Common Cause condition.

Materials.

The materials used in Experiment 1 were used in this Experiment with minor modifications. First, since the children might switch the position of the gears, two identical yellow gears were used, and the gears in the pictures were colored to match (only pictures corresponding to 1a, 1b, and 1c were used in this experiment). Because the position of the gears could now be transposed, it no longer made sense to attribute the causal power directly to the gears. Instead we taught the children that a hidden motor was under one side of the toy or that there was a motor under both sides. In each picture, we put the letter "M" (for 'motor') to indicate that the pushing gear was on the side with the motor. Thus for 1a, the M was on Gear A; for 1b, the M was on Gear B; for 1c, the M was on both gears. Because we did not attribute causal power to the gears in this experiment, we did not test children on the possibility that the gears interacted (i.e., we did not include picture 1d).

Training.

Trials began with the switch in the off position and no gears on the toy. The experimenter placed and removed each gear on the toy in turn, explaining that she could "take the gears on and off the machine". She flipped the switch on to make both gears spin simultaneously and flipped the switch off so that both gears were still. The toy was set to implement a causal chain for children in the Causal Chain conditions (particular chain, 1a or 1b, counterbalanced between children) and set to a common cause structure for children in the Common Cause condition.

The children were then told: "There are three ways this toy could work." All the children were given a choice of the two chains and the common cause structure. The children were told, "Do you know what letter this is? This is the letter 'M'. M stands for motor. Do you know what a motor is? A motor is what is inside the toy and makes the toy work. The motor is hidden inside where you cannot see it."

To illustrate each chain, the experimenter pointed to the appropriate gears and children were told: "This picture shows the motor is on this one and that this one pushes this one". For the common cause structure, children were told, "This picture shows that they both have a motor. They don't push each other. They each push themselves." Children tested singly were asked to re-explain each picture after it was introduced; children tested in dyads took turns explaining the pictures until each child successfully explained each picture.

The experimenter then shuffled the pictures, pointed to the gears on the toy and asked the children to choose the picture that showed "This one pushing this one"; or that "the wheels push themselves". All children were able to choose the correct picture. The experimenter then removed all three pictures and gave the two gears to the child. The children were told: "Go ahead and play with the toy and try to figure out how it works. You can do anything you like." The experimenter moved out of sight and supervised the children through a videocamera. If children did not stop playing themselves, the experimenter terminated the play period after five minutes.

For children tested singly, the experimenter laid out the three cards at the end of the play period and asked the children to redescribe the three cards. She then asked the child, "Can you give me the picture that shows how the toy works?" The same procedure was followed for children tested in pairs except that one child was randomly chosen to go first. The other child was escorted out of earshot to a table with art supplies and seated facing away from the gear toy. After the first child made a response, the children switched positions.

Results and Discussion

Children were coded both for the evidence they generated and the structures they chose. Children were given a score of 2 if they put a gear on only the left peg, flipped the switch on and off and also put a gear on only the right peg and flipped the switch on and off. Children were given a score of 1 if they only tested a gear in one of the two positions. Children were given a score of zero if they never tested the gears singly. (Note that all of the children also generated a wide range of actions we did not code – ranging from treating the gears like finger puppets to rolling the gears on the table like wheels.) Children's responses are shown in Table 4.

Insert Table 4 here

In the Single Causal Chain condition, half of the children generated the complete evidence. Six of the 12 children (50%) children received a score of 2; two children (17%) received a score of 1; the remaining 4 children (33%) received a score of 0. By contrast, almost all of the children who played in pairs generated the evidence. In the Dyad Causal Chain condition, 100% of the children received a score of 2. In the Dyad Common Cause condition, 10 of the 16 children (63%) received a score of 2; 4 children (25%) received a score of 1; 2 received a score of 0 (12%). Overall, children were more likely to generate complete evidence when they played in dyads than when they played alone (χ^2 (1, N = 40) = 9.41, $p < .01$).

Looking only at the children who generated the complete set of evidence, 7 of the 18 children (39%) who generated complete evidence for the causal chain (whether singly or in dyads) chose the correct causal chain. Ten children (55%) chose the common cause structure. Only one child (5%) chose the incorrect causal chain. Children were more likely to choose the correct causal chain than the incorrect chain (n = 8, $p < .05$ by binomial test) but equally likely to choose the correct chain and the common cause structure. In the Common Cause condition, 9 of the 10 children (90%) who generated the complete evidence chose the correct structure. Only one child incorrectly chose a chain. Children chose the correct structure more often than chance and more often than the incorrect structure (n = 10, $p < .05$ by binomial test). The difference between conditions was significant. When the toy was set to a chain, children chose the correct chain more than when the toy was set to a common cause; when the toy was set to a common cause, children chose the common cause more than when the toy was set to a chain (χ^2 (1, N = 38) = 3.89, $p < .05$).

These results suggest that children in this condition had a bias towards the common cause structure, possibly because the perceptually identical gears led the children to assume a causal symmetry between them. Note that Experiment 3 was less well controlled than the previous experiments. Because we did not ask children to distinguish the common cause structure (1c) from the interactive causal structure (1d), we cannot tell if children genuinely learned that the structure was a common cause or simply preferred the symmetric causal structure. Nevertheless the difference between conditions suggests that children were also learning from the evidence; although children in both conditions tended to default to the common cause structure, they were more likely to choose the correct structure (chain or common cause) when the evidence was consistent with that structure. Note also that in playing freely, children generated many interventions besides the target interventions and were exposed to variable time delays both between when they generated evidence on one gear and the other, and when they generated the evidence and were asked the test questions. Because of these factors, children’s self-generated data was much noisier than the data they observed in the earlier experiments. Nonetheless, children were often able to learn the correct causal structure from the evidence of their own interventions.

General Discussion

The results of these experiments suggest that preschoolers can use formal patterns of evidence from interventions to learn the causal structure of events and conversely, can use knowledge of causal structure to predict the outcome of novel interventions. These findings are consistent with the conditional intervention principle and support the idea that same assumptions that underlie the causal Bayes nets formalism may also be fundamental to children's causal reasoning. Moreover, in the course of their free play, children both generate evidence that could support accurate causal learning and learn from the evidence of their own interventions. This suggests that these formal inferential processes could support causal learning in real world situations and not just in the laboratory.

These results are not easily explained by other accounts of causal learning. No domain-specific features differentiated the causal structures (i.e., there was a plausible causal mechanism underlying every structure and no cues about time order or contact causality distinguished the structures). Moreover, children did not seem to be influenced by prior knowledge about how gears actually work; they were as likely to favor common cause and interactive structures as causal chains. The results are also not easily explained by other domain-general accounts of causal learning. In our study, children had to decide whether each gear played the role of a cause or an effect, and had to discriminate causal chains from common causes; in standard domain-general accounts of causal learning (e.g., Cheng, 1997; 2000; Shanks & Dickinson, 1997), the effect is established in advance and learner is asked only to discriminate among candidate causes of that effect(by their strength of association or patterns of covariation).

However, this research also raises several questions. The causal Bayes net formalism was developed to handle probabilistic data; in these experiments the data was deterministic. Although the conditional intervention principle is equally valid for both types of input, we do not know whether young children can use the conditional intervention principle to learn about causal structures when the data is stochastic. This question seems particularly critical given that, in the real world, children may more often be exposed to incomplete, noisy information than to deterministic input. Further research must look at how probabilistic evidence affects children's causal learning.

Also, although we have suggested that interventions may be central to how children think of causal relationships, evidence from interventions does not account for all of our beliefs about causes. Evidence from interventions does not seem to explain why, for instance, children assign a more important causal role to the insides than the outsides of many entities (e.g., Gelman & Wellman, 1991). Neither does the interventionist account of causation seem to explain why we believe that causal relationships are mediated by mechanisms of transmission. Further research must explore how formal inferences about causation, consistent with the interventionist account, interact with substantive, domain-specific concepts.

In their everyday life, children intervene widely on the world and see a wide range of interventions performed by others. Evidence from such interventions may give children a powerful learning mechanism for learning causal structure from data. These studies suggest that at least in generative, deterministic cases, preschoolers can use the conditional intervention principle to learn causal structures from patterns of evidence and to predict patterns of evidence from causal structure. Children seem to be able to make these inferences in the absence of differential spatiotemporal information and prior knowledge about particular causal mechanisms, and well before they can design controlled experiments themselves. Learning in very young children seems to rely on some of the same formal principles that underlie scientific discovery. In turn, these principles may help children develop intuitive theories about the world.

References

Ahn, W., Gelman, S. A., Amsterlaw, J. A., Hohenstein, J., & Kalish, C. W. (2000). Causal status effects in children's categorization. *Cognition, 76*(2), 35-43.

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; Based on a conference entitled "Cultural Knowledge and Domain Specificity," held in Ann Arbor, MI, Oct 13-16, 1990.* (pp. 169-200). New York, NY, US: Cambridge University Press.

Chen, Z. & Klahr, D., (1999) All Other Things being Equal: Children's Acquisition of the Control of Variables Strategy, *Child Development , 70*(5), 1098 - 1120.

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, US: The MIT Press.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*(2), 367-405.

Glymour, C. (2001) *The mind's arrows: Bayes nets and causal graphical models in psychology.* Cambridge, MA: MIT 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), 1-30.

Gopnik, A. & Schulz, L. E. (in press). *Causal learning: Psychology, philosophy, computation*. New York: Oxford University Press.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*(5), 620-629.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* (A. Parsons & S. Milgram, Trans.). New York: Basic Books.

Kalish, C. (1996). Causes and symptoms in preschoolers' conceptions of illness. *Child Development, 67*(4), 1647-1670.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96*(4), 674-689.

Kuhn, D., Amsel, E., & O'Laughlin, M. (1988). *The development of scientific thinking skills.* Orlando, Florida: Academic.

Kushnir, T. & Gopnik, A. (2005). Young children infer causal strength from

probabilities and interventions. *Psychological Science* 16 (9): 678-683.

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

Lagnado, D. & Sloman, S.A. (2002). Learning causal structure. In W. Gray and C. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, Fairfax, VA, 560-565.

Lehrer, R. & Schauble, L. (1998). Reasoning about structure and function: Children's conceptions of gears. *Journal of Research in Science Teaching, 31*(1) 3-25.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*(3), 265-288.

Masnick, A. M. & Klahr, D. (2003). Error matters: An initial exploration of elementary school children's understanding of experimental error. *Journal of Cognition and Development 4*(1), 67-98.

Novick, L. R., & Cheng, P. W. (in press). Assessing interactive causal influence. *Psychological Review*.

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

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

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General, 3*, 323-360.

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology, 40*(2) 162-176.

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology: Comparative & Physiological Psychology, 37*(1), 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, US: Academic Press, Inc.

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

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

Sobel, D. M. & Kushnir, T. (in press). The importance of decision making in causal learning from interventions. *Memory and Cognition*.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*(4), 605-632.

Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7*(6), 337-342.

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

Steyvers, M., Tenenbaum, J., Wagenmakers, E.J., Blum, B. (2003). Inferring Causal Networks from Observations and Interventions. *Cognitive Science, 27,* 453-489.

Tenenbaum, J. B., & Griffiths, T. L., (2003). Theory-based causal inference. In S. Becker, S. Thrun, and K. Obemayer (Eds.), *Advances in Neural Information Processing Systems 15*. Cambridge, MA: MIT Press.

Waldmann, M. R., & Hagmayer, Y. (2001). Estimating causal strength: The role of structural knowledge and processing effort. Cognition *1*, 27-58.

Watson, J. S. & Ramey, C. T. Reactions to response-contingent stimulation in early infancy. *Merrill-Palmer Quarterly 18*(3), 219-227.

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, CA, US: Jossey-Bass/Pfeiffer.

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

Table 1 *Equations associated with the gear toy.*

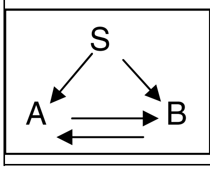
1	2	3	4
Causal graph	**Boolean equation**	**Intervention that forces A = 0**	**Intervention that forces B = 0**
S→A→B	S = 1 A = S B = A	S = 1 A = 0 B = A	S = 1 A = S B = 0
S→B→A	S = 1 A = B B = S	S = 1 A = 0 B = S	S = 1 A = B B = 0
A←S→B	S = 1 A = S B = S	S = 1 A = 0 B = S	S = 1 A = S B = 0
	S = 1 A = S * B B = S * A	S = 1 A = 0 B = S * A	S = 1 A = S * B B = 0

Table 2

Number of children choosing each picture in Experiment 1. Target responses in the Test conditions are highlighted.

	Pictures		
Group 1; n=20	**1a**	**1b**	**1d**
Trial giving evidence for structure 1a	15 (75)	1 (5)	4 (20)
Trial giving evidence for structure 1d	2 (10)	1 (5)	17 (85)
Group 2; n=20	**1a**	**1b**	**1c**
Trial giving evidence for structure 1b	3 (15)	13 (65)	4 (20)
Trial giving evidence for structure 1c	2 (10)	2 (10)	16 (80)
Group 3; n=14	**1a/1b**	**1c**	**1d**
Trial giving evidence for structure 1c	2 (14)	10 (71)	2 (14)
Trial giving evidence for structure 1d	0 (0)	2 (14)	12 (86)
Switchless Control; n=16	**1a**	**1b**	**1d**
Trial with evidence comparable to 1a	6 (37)	2 (13)	8 (50)
Trial with evidence comparable to 1d	4 (25)	4 (25)	8 (50)

Note: Percentage in parentheses (due to rounding, percentages may not sum to 100).

Table 3

Number of children making each prediction in Experiment 2. Target responses are highlighted. (Chance performance = 25%).

	Predictions			
N = 16	**Spin/Still**	**Still/Spin**	**Spin/Spin**	**Still/Still**
1a	8 (50)	3 (19)	1 (6)	4 (25)
1b	2 (12)	8 (50)	3 (19)	3 (19)
1c	4 (25)	4 (25)	7 (44)	1 (6)
1d	3 (19)	3 (19)	2 (12)	8 (50)

Note: Percentage in parentheses.

Table 4

Children's responses in Experiment 3

Children generating complete (2), partial (1), or no (0) evidence in each condition	**2**	**1**	**0**
Singles – Causal Chain (n = 12)	6 (50)	2 (17)	4 (33)
Dyads – Causal Chain (n = 12)	12 (100)	0 (0)	0 (0)
Dyads – Common Cause (n = 16)	10 (62)	4 (25)	2 (13)
	Pictures		
Children choosing each picture (of those who generated complete evidence)	**Correct chain**	**Incorrect chain**	**Common cause**
Causal Chain (n = 18)	7 (39)	1 (5)	10 (55)
Common Cause (n = 10)	0 (0)	1 (10)	90 (100)

Note: Percentage in parentheses.

Figure 1: Causal Bayes nets representing different possible causal relationships between partying (P), wine drinking (W) and insomnia (I).

1a) A causal chain

P→W→I

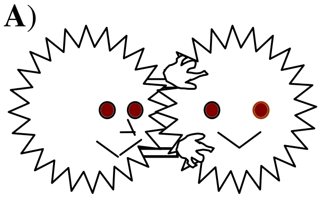
1b) A common cause model

W←P→I

Figure 2: The causal structures (boldface in the table indicates the interventions that change the outcome).

Pictures | Graph structures (S=Switch) | Patterns of evidence

A)

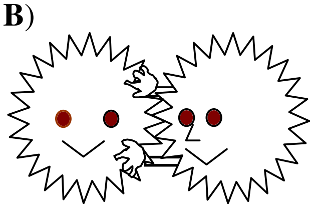


A B

S→A→B

	Interventions		**Outcome**
1	S Off	A on	B still
2	S Off	A off	B still
3	S Off	B on	A still
4	S Off	B off	A still
5	**S On**	**A on**	**B spins**
6	**S On**	**A off**	**B still**
7	S On	B on	A spins
8	S On	B off	A spins

B)

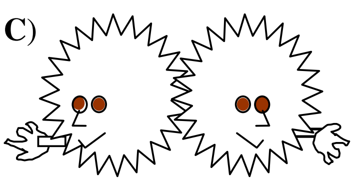


A B

S→B→A

1-4 as above			
	Interventions		**Outcome**
5	S On	A on	B spins
6	S On	A off	B spins
7	**S On**	**B on**	**A spins**
8	**S On**	**B off**	**A still**

C)

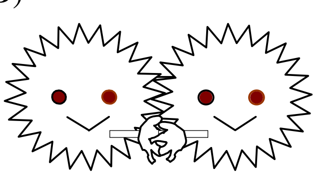


A B

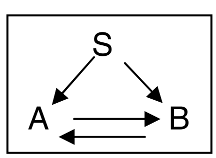
A←S→B

1-4 as above			
	Interventions		**Outcome**
5	S On	A on	B spins
6	S On	A off	B spins
7	S On	B on	A spins
8	S On	B off	A spins

D)



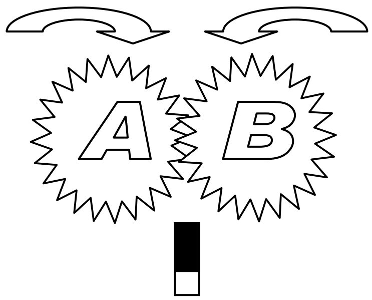
A B



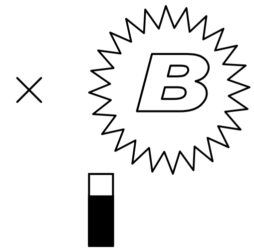
1-4 as above			
	Interventions		**Outcome**
5	**S On**	**A on**	**B spins**
6	**S On**	**A off**	**B still**
7	**S On**	**B on**	**A spins**
8	**S On**	**B off**	**A still**

Figure 3: Sample interventions and outcomes showing that A→B

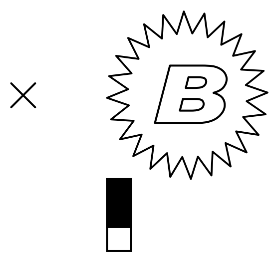
3a) Flip switch on, both gears spin



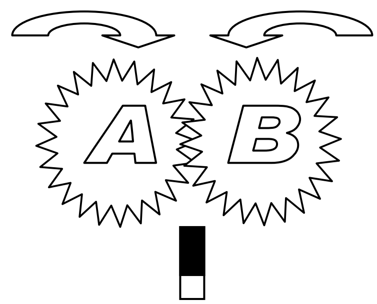
3b) Flip switch off and remove A.



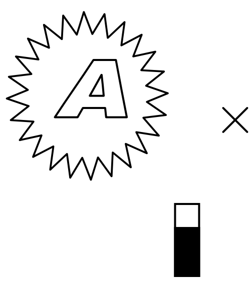
3c) Flip switch on, B doesn't spin, flip switch off



3d) Replace A, flip switch on, both spin



3e) Flip switch off and remove gear B.



3f) Flip switch on, gear A spins

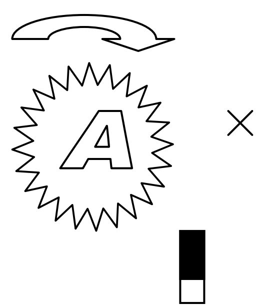


Figure 4: Schematic of the gear toy