Cognitive Science 34 (2010) 148–160

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2009.01072.x

Inferring Hidden Causal Structure

Tamar Kushnir,[a] Alison Gopnik,[b] Chris Lucas,[b] Laura Schulz[c]

[a]*Department of Human Development, Cornell University*
[b]*Department of Psychology, University of California at Berkeley*
[c]*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology*

Received 9 May 2007; received in revised form 25 July 2009; accepted 7 August 2009

Abstract

We used a new method to assess how people can infer unobserved causal structure from patterns of observed events. Participants were taught to draw causal graphs, and then shown a pattern of associations and interventions on a novel causal system. Given minimal training and no feedback, participants in Experiment 1 used causal graph notation to spontaneously draw structures containing one observed cause, one unobserved common cause, and two unobserved independent causes, depending on the pattern of associations and interventions they saw. We replicated these findings with less-informative training (Experiments 2 and 3) and a new apparatus (Experiment 3) to show that the pattern of data leads to hidden causal inferences across a range of prior constraints on causal knowledge.

Keywords: Psychology; Causal reasoning; Human experimentation; Causal inference; Causal Bayes nets; Graphical representation

1. Introduction

Our everyday theories rely on our ability to reason about unobserved causes (Gelman & Wellman, 1991; Gopnik & Meltzoff, 1997; Murphy & Medin, 1985). How might we learn about unobserved causes from observed events? We can learn about unobserved causes from descriptions of causal mechanisms (Ahn, Kalish, Medin, & Gelman, 1995) or by appeal to familiar causal principles (Gelman & Wellman, 1991; Schulz & Sommerville, 2006). However, in many cases such knowledge is incomplete or unavailable. For example, consider coming across an interesting but unfamiliar toy—a box with two sticks poking out of its surface. The sticks move up and down when operated by a hidden mechanism inside the box. Short of taking apart the toy, how might we learn what causes their movement?

Correspondence should be sent to Tamar Kushnir, Department of Human Development, Cornell University, G62B MVR, Ithaca, NY 14853. E-mail: tk397@cornell.edu

Recent research has shown that we can use statistical evidence, in particular information about both associations and interventions, to learn about unfamiliar causal relations among observed events (Gopnik, Sobel, Schulz, & Glymour, 2001; Gopnik et al., 2004; Lagnado & Sloman, 2004; Schulz, Gopnik, & Glymour, 2007; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003; Waldmann & Hagmayer, 2001). Could we also use such evidence to learn about unobserved causes—such as the causal relations governing this toy—as well?

The causal Bayes net formalism (Pearl, 2000; Spirtes, Glymour, & Scheines, 1993) provides a computational account of the conditions under which statistical evidence can lead to learning about hidden causes. The formalism defines a set of assumptions linking causal structure (directed graphs that encode causal relations) to data (conditional probabilities). Integrated into this framework is a formal notion of a causal *intervention*—a manipulation originating outside (exogenous to) the causal system that changes a variable's value in some way. Intervening on a variable temporarily eliminates the influence of that variable's parent causes within the system, as the value of the variable is completely determined by the intervention. Thus, in the causal Bayes net formalism, evidence from associations (i.e., correlations) is distinct from evidence from interventions and both are combined in a common representational framework.

Critically, this formalism is unique among psychological theories of causal inference in outlining conditions under which we can learn about unobserved causal variables and distinguish among structures containing such variables without including additional assumptions (Cheng, 1997; Shanks & Dickinson, 1987). That is not to say that other causal information—such as mechanism knowledge—cannot be incorporated. Indeed, knowledge interacts with statistical evidence to constrain the types of causal inferences that can be made in any given domain (Griffiths, Baraff, & Tenenbaum, 2004; Kushnir & Gopnik, 2007; Lagnado & Sloman, 2006; Schulz, Bonawitz, & Griffiths, 2007; Wolff, 2007). However, there is at least one scenario in which the assumptions of the formalism lead to inferences about unobserved causal structure even when no domain-specific knowledge is available. It is this pattern of associations and interventions that we consider in the following set of experiments.

To illustrate, we return to our example of the toy. Suppose that when the toy is activated, the two sticks usually move up and down together, and thus are associated. Then suppose intervening on the first stick by pulling it from above does not influence the movement of the second stick, and intervening on the second stick does not influence the movement of the first. Formally, intervening on each variable did not change the probability of the other variable. Thus, we rule out the possibility that either stick is causally responsible for the movement of the other. What possibility is left? The causal structure most consistent with the data, and the formal assumptions, is that the sticks have a hidden common cause (see Gopnik et al., 2004, for details). Therefore, the formalism predicts that we should infer the existence of a common cause, even though we do not observe it directly.

Can people learn about hidden common causes in this scenario? Previous studies on causal learning have been almost entirely restricted to relations between observed causes (e.g., Gopnik et al., 2001, 2004; Lagnado & Sloman, 2004, 2006; Steyvers et al., 2003). The technique has been to present participants with data and ask them to choose among different possible causal structures relating observed variables, with explicit instructions that there are no unobserved causes. One study (Kushnir, Gopnik, Schulz, & Danks, 2003) used this technique to show that people can choose the correct causal structure containing unobserved causes. However, we do not yet know whether people can posit hidden causal structure spontaneously, without being given a set of choices.

The current study presents a new method designed to examine whether people can make spontaneous inferences about unobserved causal structure independently of specific knowledge of causal mechanisms. Similar to previous studies (Lagnado & Sloman, 2006; Sobel & Kushnir, 2006), we showed participants how to draw arrows between variables to indicate causal relations. However, whereas in previous research variables were prespecified, here we purposefully left variable selection up to participants. Importantly, participants were told they could include hidden causal variables when necessary.

After a brief training on drawing causal graphs, we showed participants a novel causal system (a toy) with two observed variables. Then they saw four different patterns of evidence involving associations and interventions. Each pattern was consistent with a different causal structure connecting the observed variables, some of which necessarily included hidden causal variables. We hypothesized that, consistent with the predictions of the causal Bayes net formalism, participants would spontaneously infer hidden common causes when they observed variables that are generally associated but are independent under intervention. Moreover, we predicted that participants could make inferences about hidden causal structure without prior constraints on the possible causal structures they could choose to draw (Experiment 2) and independently of their knowledge of particular physical causal mechanisms (Experiment 3).

2. Experiment 1

Experiment 1 introduces a new method—drawing causal graphs—to test people's inferences about hidden causal structure. We provided no feedback to participants about their answers at any time during the training or test conditions. The four conditions in this experiment all involved a ''stick-ball machine'' and were based on those in Kushnir et al. (2003) and are described in detail below.

2.1. Method

2.1.1. Participants

Fourty-six undergraduates were recruited from the research participation pool at a large university. They were tested in four roughly equal groups (11–12 participants per group).

2.1.2. Materials

The stick-ball machine (Fig. 1) was a $3' \times 1' \times 1'$ wooden box with two holes at the top and an open back hidden from participants. Two balls attached to sticks were placed in the holes. In *association trials*, the sticks were moved up and down (either together or separately) from behind—thus participants observed associations between the two sticks'

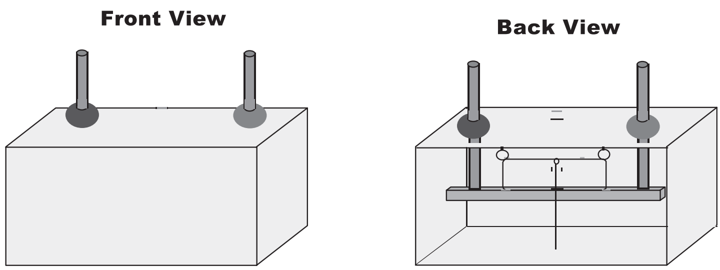


Fig. 1. A drawing of the stick-ball machine apparatus from Experiments 1 and 2. The sticks could be connected and moved by a common beam (shown) or could be disconnected from the beam and moved separately (for the conditions requiring independent movement). Note that participants were not given any detailed description of the inner workings of the apparatus.

movements. These associations might have been caused by a mechanism that involved one stick moving the other, or involved two sticks being moved independently, or involved a common cause moving both sticks together—prior mechanism knowledge could not distinguish these possibilities. In *intervention trials*, participants saw an experimenter visibly pull one of the sticks up and observed the effect on the other stick.

2.1.3. Procedure

2.1.3.1. Graph training phase: We trained participants to draw causal graphs as containing circles, representing ''things,'' and arrows, representing causal direction. Participants practiced with five examples in which large colored blocks hit each other domino-style: (1) domino A knocked over domino B, (2) domino B knocked over domino A, (3) domino C knocked over dominoes A and B, and (4) domino C knocked over domino A and domino D knocked over domino B. In a final example (5), one block was on the table and fell spontaneously. Participants were told that if they did not see the cause of an event, they could use an ''h'' to refer to a hidden cause. No feedback of any sort was given.

Test phase. One experimenter narrated the task and performed interventions while the other operated the machine. Participants were given the task to draw the mechanism that made the stick-balls move in each condition, and they were told that it could change from condition to condition. No training on interventions—verbal or visual—was included.

Each condition contained two new stick-balls of different colors. Familiarization was always first, followed by the test conditions counterbalanced (by group of participants) in a Latin square design. The following descriptions of the conditions are summaries. In the actual experiment, the types of movement (interventions and associations) within each condition were intermixed. The interventions were counterbalanced by side so that no stick (right or left) was always intervened on first.

2.1.3.2. Familiarization: This established that the causal relations were probabilistic rather than deterministic. Participants were given the causal structure: Stick A caused stick B to move. This was then demonstrated by showing both sticks moving together four times and stick A moving alone twice, with no interventions. Griffiths et al. (2004) previously showed that this pattern of movements on the stick-ball machine leads participants to expect causes to be 75% effective.

1. *One observed-cause condition.* The narrator intervened on stick A six times. Four of those times, both sticks A and B moved. The remaining two times stick A moved and stick B did not move. Thus, A and B were highly correlated due to the intervention on A. Learners should conclude that A causes B.
2. *Common unobserved-cause condition.* In association trials the stick-balls moved together four times. The narrator intervened on stick A twice and each time stick B did not move. The narrator intervened on stick B twice and each time stick A did not move. So A and B were strongly correlated in the association trials, but independent under interventions. Learners should therefore infer the presence of a hidden common cause of A and B.
3. *Independent unobserved-causes condition.* In association trials, the stick-balls moved separately twice and together once. The narrator intervened on stick A twice and each time stick B did not move. The narrator intervened on stick B twice and each time stick A did not move. Note that the interventions were the same as in the common unobserved-cause condition. However, the correlation between A and B in the association trials was weak. Learners should conclude that there are independent causes for A and B.
4. *Pointing control condition.* In association trials, the stick-balls moved together four times. The narrator pointed at stick A twice as it moved alone. The narrator pointed at stick B twice as it moved alone. Pointing began slightly after the movement (to rule it out as a cause). This condition was included to rule out two potential alternative hypotheses for performance in the hidden cause condition: associative learning and salience. First, the overall movements of the stick-balls were identical to the common unobserved-cause condition; in each case the movements occurred together 50% of the time and separately 50% of the time. Second, pointing was included to ensure that participants understood that only the appropriate pulling gesture could be treated as an exogenous intervention; the mere presence (and salience) of a hand gesture should not be sufficient. We predict that, when the independent movements of the sticks were accompanied by pointing (rather than pulling) participants should be uncertain about the underlying causal structure.

After each trial participants were told to draw a graph that corresponded to the causal structure.

Coding. Responses were coded as either (1) one cause, (2) hidden common cause, (3) hidden independent causes, or (4) other. Eighty percent of the responses were coded as 1, 2, or 3. The additional 20% were coded as ''other'' because they either contained multiple structures (either side by side or superimposed on the same graph, 6%), had a bidirected edge (could have been interpreted as cyclic, conjunctive, or as containing a common cause, 6.5%), or were otherwise unclear (7.6%). A small number of participants[1] that referred to a person (or hand) causing the machine's activation, but made responses consistent with graphs of types 1, 2, or 3 were coded as falling into the appropriate response category. An additional check was made to make sure that the ''one-cause'' response in the one-cause trial was in the appropriate direction, which was true for all cases. Coding was performed by the first author and by an independent coder who was blind to the research question and trial order. Raters were in agreement on 98% of the responses. Importantly, raters were in 100% agreement about hidden common-cause responses.

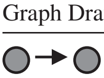
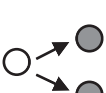
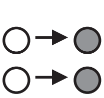
2.2. Results and discussion

All of the participants drew the correct causal graphs in the training phase. Moreover, in the test phase, the responses (described below) were consistent with those obtained from asking participants to choose among a set of given hidden causal structures (Kushnir et al., 2003). We therefore believe this to be a promising method for assessing causal inferences in future research.

Table 1 shows the percentage of participants drawing each type of graph in each condition. The majority response in each of the first three conditions was the most likely graph given the pattern of associations and interventions. In the one observed-cause condition, 71.7% drew the correct one-case structure.[2] In the common unobserved-cause condition, 58.7% drew a hidden common-cause graph. In the independent unobserved-causes condition, 52.2% drew the independent-causes graph.

Importantly, within-subjects comparisons showed that individual participants drew the appropriate graph in each condition. All comparisons were made using McNemar's tests (two-tailed binomial tests on the change proportions). All tests were significant at an alpha level of .001 unless otherwise noted. More participants drew the one-cause graph in condition 1 but not condition 2 (31 vs. 0) or condition 3 (33 vs. 0). Moreover, more participants drew a hidden common-cause graph in condition 2 but not condition 1 (24 vs. 0) or condi-

Table 1
Percentage of responses in each of the test conditions in Experiment 1

	Condition			
Graph Drawn	One Observed	Common Unobserved	Independent Unobserved	Pointing Control
	**71.7**[a]	4.3	0	6.5
	6.5	**58.7**	28.3	39.1
	8.7	17.4	**52.2**	26.1
Other	13	19.6	19.6	28.3

Note. The modal response is indicated in bold.
[a]This was the correct one-cause answer only; no participant drew this graph in the other direction.

tion 3 (16 vs. 2). Finally, more participants drew the independent-causes graph in condition 3 but not condition 1 (20 vs. 0) or condition 2 (16 vs. 0).

In the pointing control (condition 4), we predicted that participants should be uncertain about the correct structure, and the results confirm this prediction. Participants were unlikely to choose the one-cause graph (6.5%); but there were no differences between the common-cause and independent-causes responses, $\chi^2(1, N = 30) = 1.2$, ns. Importantly, McNemar's tests showed that participants made significantly more hidden common-cause responses in condition 2 (12 vs. 3, $p < .05$) than in this condition and made significantly more independent-causes responses in condition 3 than in this condition (12 vs. 0).

3. Experiment 2

Given minimal training and no feedback, participants in Experiment 1 used causal graph notation to spontaneously draw structures containing one unobserved common cause, two unobserved independent causes, or one observed cause depending on the pattern of associations and interventions. However, the training included observed causal structures corresponding to exactly the hidden structures in the test condition—the participants saw common-cause and independent-cause structures in both cases. This may have potentially biased the participants to favor those responses. In Experiment 2, we trained participants on different graphs to eliminate this possibility.

3.1. Method

3.1.1. Participants

Participants were 44 undergraduates recruited from the research participation pool at a large research university. They were tested in four equal groups.

3.1.2. Materials

The materials were the same as in Experiment 1.

3.1.3. Procedure

The graph training was identical to the training in Experiment 1, except that scenarios 3 and 4 (common cause and independent causes) were replaced by (3) *Common effect*: dominoes A and B knocked over domino C, and (4) *Chain*: domino C knocked over domino A and domino A knocked over domino B.

Participants then saw the same familiarization condition and four test conditions, counterbalanced as in Experiment 1.

Graphs drawn were coded by the criteria used in Experiment 1. There were no errors in the training phase. Eighty-six percent of the graphs in the test phase were coded as 1, 2, or 3. All but four of the ''one-cause'' response in the once-cause trial were in the appropriate direction. Those four were recoded as ''other'' for the remainder of the analysis. The first author and a blind coder were in agreement on 94% of the responses. Both coders were in 97% agreement about hidden common-cause responses.

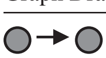
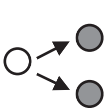
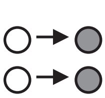
3.2. Results and discussion

The results replicate those of Experiment 1; participants drew the appropriate graphs without being explicitly trained on the structures beforehand. Table 2 shows the percentage of participants drawing each type of graph in each condition. In the one observed-cause condition, 88.6% drew the correct one-cause graph. In the common unobserved-cause condition, 47.7% drew a hidden common-cause graph. In the independent unobserved-causes condition, 50% drew an independent-causes structure.

As in Experiment 1, individual participants drew the appropriate graphs in each condition. All within-subject comparisons were analyzed by one-tailed McNemar's tests based on a hypothesized replication of the direction of effect. Results were significant at an alpha level of .001 unless otherwise noted. More participants drew the one-cause graph in condition 1 but not condition 2 (37 vs. 2) or condition 3 (37 vs. 1). Moreover, more participants drew a hidden common-cause graph in condition 2 but not condition 1 (21 vs. 0) or condition 3 (11 vs. 3; $p < .05$). Finally, more participants drew the independent-causes graph in condition 3 but not condition 1 (22 vs. 0) or condition 2 (15 vs. 4; $p < .01$).

In the pointing control, roughly equal numbers of participants drew graphs 1, 2, and 3, $\chi^2(2, N = 34) = 0.77$, ns. Critically, comparisons of the hidden causal structures between condition 4 and conditions 2 and 3 showed the same effect as in Experiment 1—more participants drew the hidden cause graph in condition 2 (11 vs. 3; $p < .05$), and more drew the independent cause graph in condition 3 (12 vs. 2, $p < .01$) than in the pointing control. Thus, participants inferred the correct structure when the pattern of data included the appropriate

Table 2
Percentage of responses in each of the test conditions in Experiment 2

	Condition			
Graph Drawn	One Observed	Common Unobserved	Independent Unobserved	Pointing Control
	**88.6**[a]	9.1	6.8	20.5
	0	**47.7**	29.5	29.5
	0	25.0	**50.0**	27.3
Other	11.4[b]	18.2	13.6	22.7

Notes. The modal response is indicated in bold.
[a]This was the correct one-cause answer only.
[b]This included the one-cause response drawn in the wrong direction.

associations and interventions, although they were not given any explicit training to draw either common-cause graphs or independent-causes graphs.

4. Experiment 3

The results of Experiment 2 suggest that participants not only inferred the correct structure from the data but were also able to generate the hypothesis space of structures themselves. However, it is quite possible that the physical constraints of the stick-ball machine itself were driving participants' inferences about the possible causal structures. Indeed, Griffiths et al. (2004) showed computationally how a simple physical theory can constrain the hypothesis space of causal structures governing the operation of the stick-ball machine. It is not known whether the pattern of evidence we present here can lead to inferences about hidden causal structure across different naive physical theories.

To examine this, we replicated Experiment 2 using a different causal mechanism—one that was electronic rather than mechanical, and therefore governed by different physical constraints. Specifically, our intuitions about electronic mechanisms are that they are more complex than mechanical ones (Koslowski, 1996). A more complex physical theory would lead to a larger hypothesis space of causal structures connecting the observed variables (Griffiths et al., 2004). Nonetheless, we predict that the general pattern of results would replicate with this new device.

4.1. Method

4.1.1. Participants

Participants were 54 undergraduates at a large university. Eleven additional participants were excluded because they drew at least one incorrect graph in the training phase or familiarization condition (when the answer was known). Participants were tested in 11 small groups.

4.1.2. Materials

The apparatus consisted of two electronic devices: $5'' \times 7'' \times 3''$ boxes with orange panels on top (Fig. 2). When activated, the devices' orange panels glowed and they played a short melody. In association trials, one or both of the devices were activated by hidden switches controlled by a confederate. In intervention trials, the experimenter activated one device by pressing on its orange panel, and the other device activated simultaneously or not at all.

4.1.3. Procedure

Participants saw the same graph training, familiarization condition, and four test conditions as in Experiment 2, counterbalanced in one of four orderings as before.

Coding criteria followed the previous experiments. Sixty-nine percent of graphs were coded as 1, 2, or 3. The responses included a small number of graphs (5%), not present in

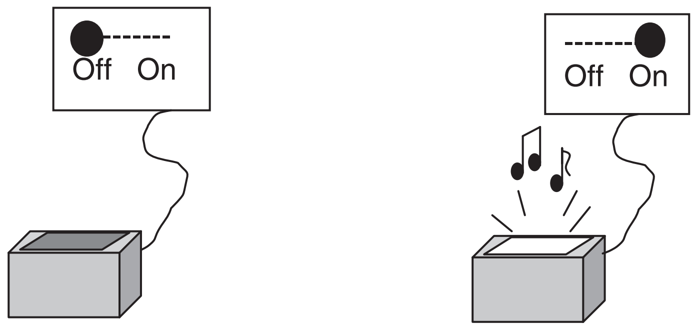


Fig. 2. A drawing of the electrical devices used in Experiment 3 (switches were hidden from view).

Experiments 1 and 2, depicting independence between observed variables without hidden causes. These were also coded as response type 3. Interestingly, these graphs bear a strong physical resemblance to the way the toy actually looked (two toys, no wires or other connections between them). The statistical comparisons below are equivalent with or without these responses included. Coding was performed by the third author and a hypothesis and condition blind coder. Agreement was 94.2%. Coders were in 93.8% agreement about hidden common-cause responses.

4.2. Results and discussion

Overall, the results (shown in Table 3) show the same pattern as in Experiments 1 and 2. In the common unobserved-cause condition, 48.1% of participants drew the correct hidden common-cause graph. In the independent unobserved-causes condition, 33.3% drew the correct independent-causes graph. In the one observed-cause condition, 63% drew the correct one-cause graph.

Table 3
Percentage of responses in each of the test conditions in Experiment 3

	Condition			
Graph Drawn	One Observed	Common Unobserved	Independent Unobserved	Pointing Control
● → ●	**63**[a]	13	11.1	25.9
○ ↗ ● ↘ ●	7.4	**48.1**	24.1	24.1
○ → ● ○ → ●	5.6	7.4	**33.3**	13
Other	24	31.5	31.5	37

Note. The modal response is indicated in bold.
[a]This was the correct one-cause answer only; no participant drew this graph in the other direction.

As before, individual participants drew the appropriate graphs in each condition. All comparisons in this section were made using McNemar's tests (two-tailed) and were significant at an alpha level of .001 unless otherwise noted. Participants were more likely to draw the hidden common-cause graph in condition 1 and not condition 2 than vice-versa. A similar pattern held for in condition 2 (15 vs. 2, $p < .01$) and condition 3 (23 vs. 1). Similarly, they were more likely to draw the hidden independent-causes graph in condition 2 than in condition 1 (15 vs. 1, $p = .001$) or condition 3 (16 vs. 1). They were also more likely to draw the one-cause graph in condition 3 only than in condition 1 (28 vs. 1) or condition 2 (28 vs. 0).

In the pointing control, the distribution was uniform across all four response categories, $\chi^2(3, N = 54) = 3.04$, ns, replicating the findings from Experiment 2. Once again, significantly more participants drew the one-cause graph in condition 1 (27 vs. 7), hidden-cause graph in condition 2 (15 vs. 2, $p < .01$), and independent-causes graph in condition 3 (13 vs. 2, $p < .01$) than in the pointing control condition.

As expected, in Experiment 3, there were a slightly larger number of responses (31%) coded as ''other'' (as compared to only 16% in Experiment 2 and only 20% in Experiment 1). This is consistent with our prediction that the electronic device elicits a larger hypothesis space of causal structures. Despite this difference, the majority of participants used the evidence to spontaneously generate the appropriate structure in each condition.

5. General discussion

This study demonstrates that, given minimal training and no prior constraints on responses, people can spontaneously draw causal graphs containing hidden variables. In all three experiments, participants drew the appropriate graphs containing hidden common causes when they observed that two events were associated but the association was not preserved under interventions. In addition, participants correctly inferred independent causes when they observed the same pattern of interventions without an observed association (condition 3), and they were appropriately agnostic about the causal structure underlying the same pattern of movements and associations without interventions (condition 4).

This study also shows that people can infer hidden common causes from the same pattern of evidence across different causal mechanisms. Previous research has shown that human causal reasoning relies heavily on knowledge of mechanisms (Ahn et al., 1995; Koslowski, 1996), but what if mechanisms are unknown? There is considerable evidence that both adults and children have relatively shallow knowledge of causal mechanisms (Keil, 2003). One way to deepen our understanding of new mechanisms may be to observe patterns of associations and interventions similar to those in this study.

The Causal Bayes net formalism defines conditions under which hidden causes can be discovered independently of domain-specific knowledge. Human causal learning, however, involves the interaction of domain-general and domain-specific causal knowledge, and recent studies have begun to examine the scope of this interaction (see Tenenbaum & Griffiths, 2007). Moreover, other domain-general assumptions likely play a role in hidden causal discovery. There is evidence that adults and young children assume causal determinism, that is, they believe that probabilistic relationships between causes and effects can be explained by appealing to unobserved causes (Luhmann & Ahn, 2003; Schulz & Sommerville, 2006). Similarly, adults and children seem to assume that there are unobserved mechanisms that transmit causal power from causes to their effects (Ahn et al., 1995; Shultz, 1982). A combination of these processes may best explain how we are able to infer complex unobserved causal structure from observed events. In turn, this kind of learning may play an important role in the development of our intuitive theories.

Notes

1. Seven percent across all three studies.
2. Griffiths et al. (2004) computationally showed that expected response distributions in this study are probabilistic rather than all-or-none and that alternative responses reflect individual differences in the willingness to entertain other plausible (but less probable) hypotheses. Our data are consistent with this interpretation.

References

Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *57*(3), 299–352.

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

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

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 1–31.

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

Gopnik, A., Sobel, D. M., Schulz, L., & 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.

Griffiths, T. L., Baraff, E. R., & Tenenbaum, J. B. (2004). Using physical theories to infer hidden causal structure. In K.D. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, *18*(5–6), 663–692.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.

Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology*, *44*, 186–196.

Kushnir, T., Gopnik, A., Schulz, L. E., & Danks, D. (2003). Inferring hidden causes. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th annual meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Lagnado, D. A., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 856–876.

Lagnado, D. A., & Sloman, S. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 451–460.

Luhmann, C. C., & Ahn, W. (2003). Evaluating the causal role of unobserved variables. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th annual conference of the Cognitive Science Society* (pp. 734–739). Mahwah, NJ: Erlbaum.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289–316.

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

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

Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, *10*(3), 322–332.

Schulz, L. E., & Sommerville, J. (2006). God does not play dice: Causal determinism and children's inferences about unobserved causes. *Child Development*, *77*(2), 427–442.

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.

Sobel, D. M., & Kushnir, T. (2006). The importance of decision-making in causal learning from interventions. *Memory & Cognition*, *34*(2), 411–419.

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

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453–489.

Tenenbaum, J., Griffiths, T., & Niyogi, S. (2007). Intuitive theories as grammars for causal inference. In A. Gopnik & L. E. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 301–322). New York: Oxford University Press.

Waldmann, M. R., & Hagmayer, Y. (2001). Estimating causal strength: The role of structural knowledge and processing effort. *Cognition*, *82*, 27–58.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General*, *136*, 82–111.