Mechanisms of theory formation in young children

Alison Gopnik and Laura Schulz

Department of Psychology, University of California at Berkeley, Berkeley, CA, 94720, USA

Research suggests that by the age of five, children have extensive causal knowledge, in the form of intuitive theories. The crucial question for developmental cognitive science is *how* young children are able to learn causal structure from evidence. Recently, researchers in computer science and statistics have developed representations (causal Bayes nets) and learning algorithms to infer causal structure from evidence. Here we explore evidence suggesting that infants and children have the prerequisites for making causal inferences consistent with causal Bayes net learning algorithms. Specifically, we look at infants and children's ability to learn from evidence in the form of conditional probabilities, interventions and combinations of the two.

Over the past 30 years we have discovered an enormous amount about what children know and when they know it. In particular, young children, and even infants, seem to have intuitive theories of the physical, biological and psychological world (for recent reviews see [1–3]). These theories, like scientific theories, are complex, coherent, abstract representations of the causal structure of the world. Even the youngest preschoolers can use these intuitive theories to make causal predictions, provide causal explanations, and reason about causation counterfactually [4–7]. Moreover, both studies of natural variation in relevant experiences, and explicit training studies, demonstrate that children's intuitive theories change in response to evidence [8–11].

But the real question for developmental cognitive science is not so much what children know and when they know it, but how children's theories develop and change and why children's theories converge towards accurate descriptions of the world. It is all very well to suggest that children's learning mechanisms are analogous to scientific theory-formation. However, what we would really like is a more precise specification of the mechanisms that underlie learning in both scientists and children.

One such candidate learning mechanism has recently attracted considerable interest within the fields of computer science, philosophy and psychology. The causal Bayes net account of causal knowledge and learning provides computational learning procedures that allow abstract, coherent, structured representations to be derived from patterns of evidence, given certain assumptions [12–15]. One advantage of this formal learning account is that it specifies, with some precision, the kinds of abilities that must be in place in order for learning to occur. We will give an overview of the causal Bayes net formalism and then outline recent research regarding two foundational types of abilities that would support causal learning within this formal account. Some aspects of these abilities have already been investigated empirically, but we will also point to crucial questions that have yet to be explored.

Causal Bayes nets

Causal directed graphical models, or causal Bayes nets, have been developed in the philosophy of science and statistical literature over the last 15 years [12–15]. The models provide a formal account of a kind of inductive inference that is particularly important in scientific theory-formation. Scientists infer causal structure by observing the patterns of conditional probability among events (as in statistical analysis) by examining the consequences of interventions (as in experiments) or, usually, by combining the two types of evidence. Causal Bayes nets provide a mathematical account of these inferences and so a kind of inductive causal logic.

Causal relations are represented by directed acyclic graphs. The graphs consist of variables, representing types of events or states of the world, and directed edges (arrows) representing the causal relations between those variables (see Figure 1). The structure of a causal graph constrains the probability of the variables in that graph. In particular, it constrains the CONDITIONAL INDEPENDENCIES among those variables (see Glossary). These constraints can be captured by a single formal assumption: the CAUSAL MARKOV ASSUMPTION. The causal Markov assumption
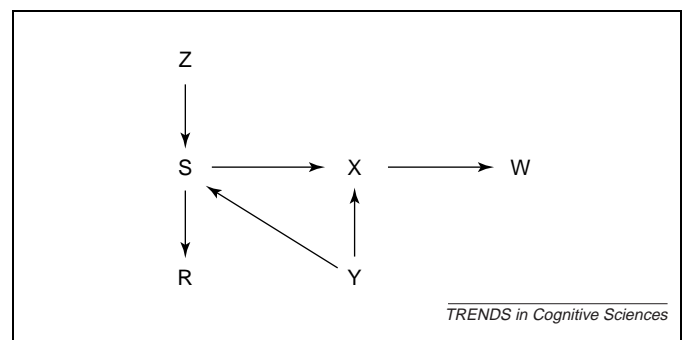

Figure 1. A causal Bayes net. R, S, W, X, Y, Z represent variables and the arrows represent causal relations between those variables.

Corresponding author: Alison Gopnik (gopnik@socrates.berkeley.edu).
Available online 8 July 2004

specifies that, given a particular causal structure, only some patterns of conditional independence will occur among the variables. Therefore, we can use knowledge of the causal graph to predict the patterns of conditional probability.

The constraints also allow us to determine what will happen when we intervene from outside to change the value of a particular variable. When two variables are genuinely related in a causal way then, holding other variables constant, intervening to change the value of one variable should change the value of the other. Indeed, philosophers have recently argued that this is just what it means for two variables to be causally related [16,17]. If we assume a particular formal definition of intervention (the INTERVENTION ASSUMPTION), we can use causal Bayes nets to predict the effects of interventions on a causal structure. A central aspect of causal Bayes nets, indeed the thing that makes them causal, is that they allow us to freely go back and forth from evidence derived from observations to inferences about interventions and vice-versa.

We can also use the formalism to work backwards and learn the causal graph from patterns of conditional probability and intervention. This type of learning requires a third assumption: the FAITHFULNESS ASSUMPTION. Given the faithfulness assumption, it is possible to infer complex causal structure from patterns of conditional dependence and independence and intervention. In some cases, it is also possible to accurately infer the existence and even the structure of new unobserved variables that are common causes of the observed variables [18,19]. Computationally tractable learning algorithms have been designed to accomplish these tasks and have been extensively applied in a range of disciplines [e.g.,20,21].

Recently, several investigators have suggested that adults' causal knowledge might involve implicit forms of Bayes nets representations and learning algorithms [22–27]. However, adults have extensive experience and often, explicit tuition in causal inference. If young children could use versions of Bayes nets assumptions and computations they would have a powerful tool for making

causal inferences. They might, at least in principle, use such methods to uncover the kind of causal structure involved in everyday intuitive theories. [28,29] However, learning of the sort represented by the causal Bayes net formalism requires: (i) the ability to learn from conditional probabilities, (ii) the ability to learn from interventions, and (iii) the ability to combine these two types of learning. Is there any evidence that young children have these prerequisite abilities?

Learning from conditional probabilities
The basic data for Bayes net inferences are judgments about the conditional independence of variables, judgments that require computing the conditional probabilities of values of those variables. There has recently been a great deal of work suggesting that, given non-causal data, such probabilities are computed spontaneously even by infants [30]. One such finding showed that eight-month-old infants could calculate the conditional probabilities of linguistic syllables in an artificial language [31]. Since then the experiments have been replicated with non-linguistic tones [32], with simultaneous visual stimuli [33], and with temporal sequences of visual stimuli [34]. These findings suggest that conditional probability information is available to infants and may be translated into more abstract representations.

There are still, however, many unanswered questions. Previous experiments have pitted conditional probabilities of 1 against those of less than one (usually 0.33), and shown that infants can distinguish these levels of probability. We do not know if infants can discriminate among finer degrees of conditional probability. Moreover, we do not know if infants can calculate CONDITIONAL DEPENDENCE and INDEPENDENCE, that is, whether they can tell that one stimulus is dependent on another only conditional on some other stimulus (a kind of conditional conditional probability). Finally, we do not know whether infants' ability to track the conditional probability of non-causal stimuli in these domains extends to an ability to track the conditional probability of candidate causes and effects. However, studies answering these questions should be feasible with the existing techniques.

We do know more about conditional probability judgments in young children. Clearly, young children cannot explicitly and consciously relate conditional probability to causation. However, we can show children novel causal relations among novel types of events, for example, by presenting them with a newly-invented machine. We give children information about the conditional probabilities of those events and see what causal conclusions they draw.

Two-and-a-half-year-olds can discriminate conditional independence and dependence, that is, conditional conditional probabilities, even with controls for frequency, and can use that information to make judgments about causation [35]. In these experiments children saw various combinations of objects placed on a machine, which did or did not light up. The children were told that 'blickets make the machine go' and were asked to identify which objects were blickets. For example, children saw the sequence of events depicted in Figure 2a, and the control sequence depicted in Figure 2b. In Figure 2a the effect *E* (the
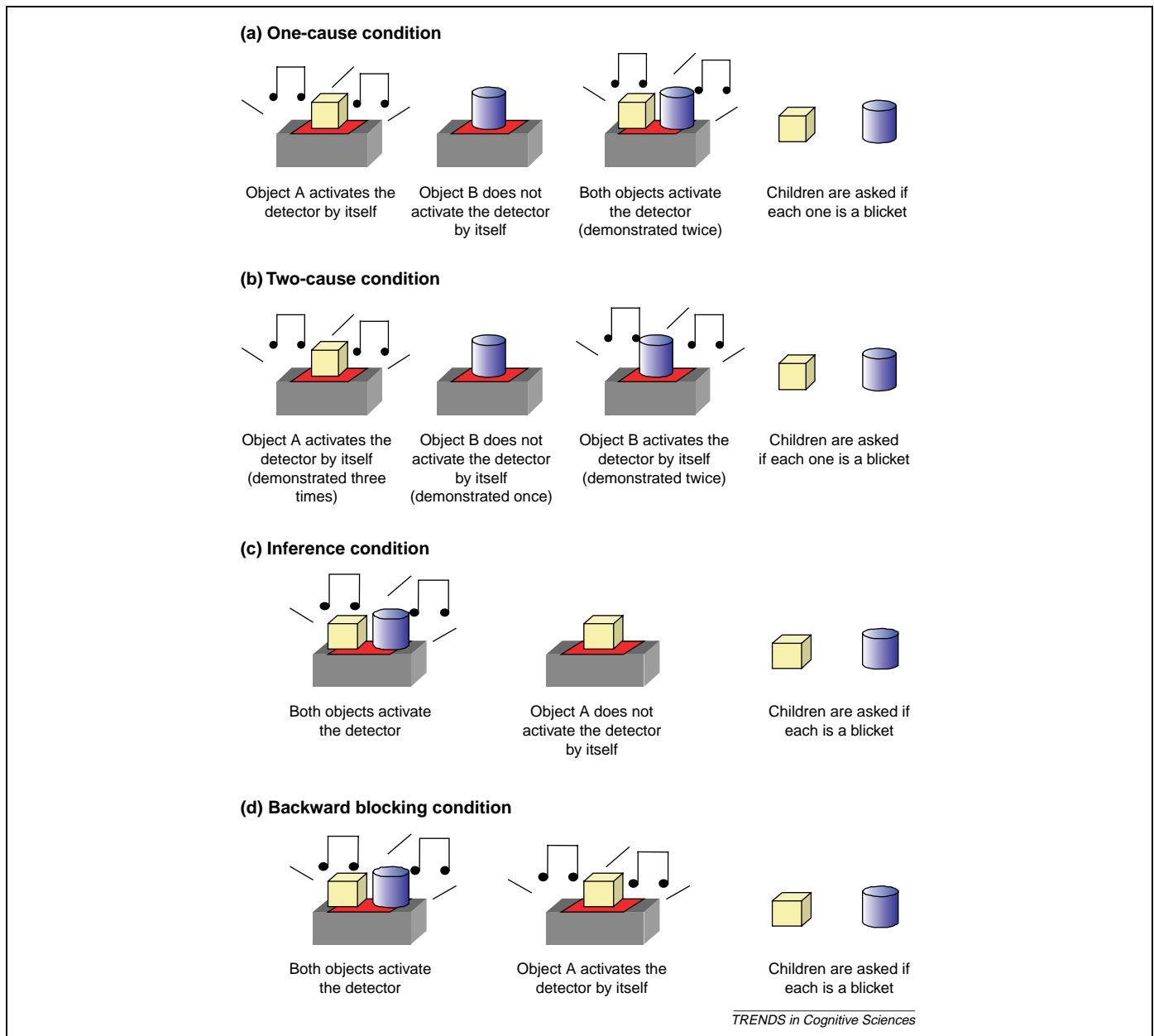
Figure 2. Screening-off and backwards blocking. In the screening-off procedure [35], children are presented with two conditions: **(a)** In the one-cause condition, only object A causes the machine to go; **(b)** In the control two-cause condition, both A and B cause the the machine to go. In the backwards-blocking procedure [36], there are also two conditions: **(c)** In the inference condition only B makes the machine go; **(d)** In backwards blocking, A makes the machine go, and B may or may not make the machine go. (See text for results.)

detector lighting up) is correlated with both object A and object B. However, *E* is independent in probability of B conditional on A, but *E* remains dependent on A conditional on B. In Figure 2b each block activates the detector the same number of times as in Figure 2a but the conditional independence patterns are the same for A and B. Children consistently choose A rather than B as the blicket, in the first condition, and choose equally between the two blocks in the second condition. Assuming that the causal relations are deterministic, generative and non-interactive, a Bayes net account would generate a similar conclusion.

Moreover, in similar experiments, four-year-old children used principles of Bayesian inference to combine prior probability information with information about the

conditional probability of events. [36]. For example, suppose children see the sequence of events in Figures 2c and 2d. On a Bayes net account, the causal structure of 2c is clear: A does not cause the effect and B does, and the children also say this. However, the causal structure of 2d is ambiguous, it could be that A and B both make the detector go, but it is also possible that only A does. Indeed, children give both types of responses. However, we can increase the prior probability of the 'A only' structure by telling the children beforehand that almost none of the blocks are blickets. Children who are told that blickets are rare are more likely to choose the 'A only' structure – that is to say that A is a blicket but B is not.

Four-year-olds can also perform even more complex kinds of reasoning about conditional dependencies, and
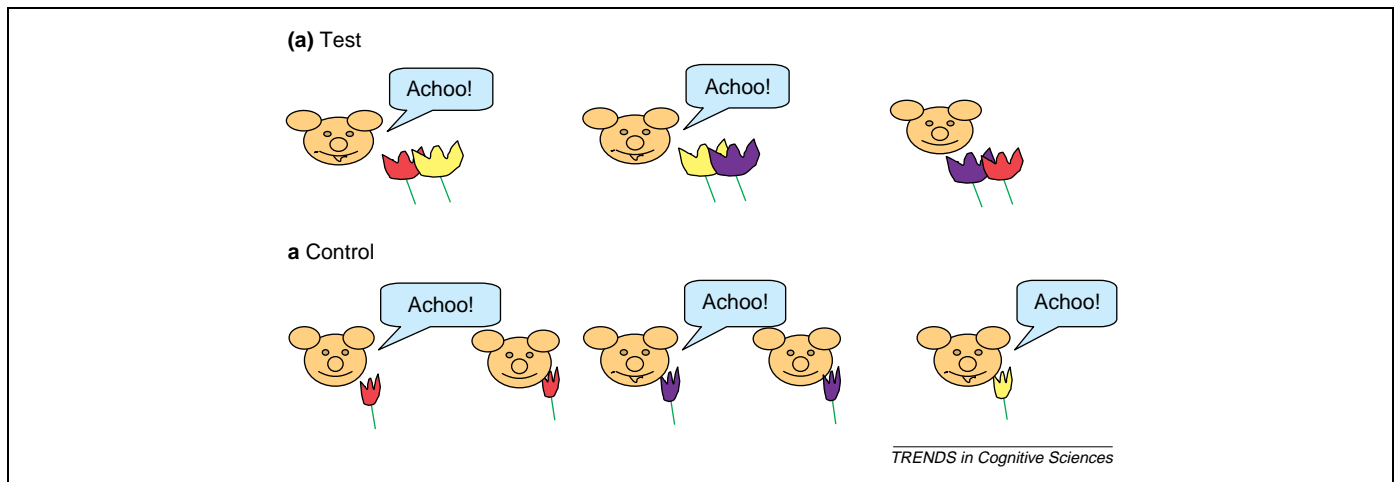
Figure 3. Screening-off in a biological task [37]. **(a)** Test condition: children see that the red and yellow flowers together make Monkey sneeze and that the blue and yellow flowers together make Monkey sneeze, but that the red and blue flowers together do not make Monkey sneeze. **(b)** Control condition: children see identical frequency information but each flower is presented singly; the red and blue flower each make Monkey sneeze half the time; the yellow flower makes Monkey sneeze all the time. In each condition, children are asked which flower makes the Monkey sneeze. Children correctly choose the yellow flower in the test condition but choose at chance in the frequency control condition.

they do so in many domains, biological and psychological as well as physical. In one experiment children were shown a monkey puppet and various combinations of flowers in a vase (see Figure 3). They were told that some flowers made the monkey sneeze and others didn't. Then they were shown the following sequence of events: Flowers A and B together made monkey sneeze. Flowers A and C together made monkey sneeze. Flowers B and C together did not make monkey sneeze. Children correctly concluded that A would make the monkey sneeze by itself, but B and C would not [37]. In a frequency control condition, in which flowers B and C made monkey sneeze half the time and flower C all the time, children chose each of the three flowers equally often.

Learning from interventions

Conditional probability is one basic type of evidence for causation. The other basic type of evidence involves understanding interventions and their consequences. The technical definition of the INTERVENTION ASSUMPTION might look formidable but it actually maps well onto our everyday intuitions about intentional goal-directed human actions. We assume that such actions are the result of our freely willed mental intentions, and so unaffected by the variables they act on (Clause 1). Clause 2 is basic to understanding goal-directed action. When actions are genuinely goal-directed we can tell whether our actions are effective: that is whether they determine the state of the variables we act upon, and we modify the actions if they are not. Clause 3 is essential to understanding means–ends relations. When we act on means to gain an end we assume that our actions influenced other variables (our ends) through, and only through, the influence on the acted-upon variable (the means).

Moreover, we assume that these features of our own interventions are shared by the interventions of others. This is an important assumption because it greatly increases our opportunities for learning about causal structure – we learn not only from our own actions but also from the actions of others.

Several features of this understanding of intervention appear to be in place at a very early age. In terms of Clause 2, infants seem to 'parse' sequences of human actions into meaningful goal-directed units [38,39]. By around seven months of age, infants understand at least some particular goals of human action and understand that goal-directed actions should be understood differently than interactions between objects [40–42]. For instance, if infants see a hand reach several times towards a particular object and the location of the object is changed, infants look longer when the hand reaches to a new object in the familiar location (i.e. the goal changes) than the familiar object in a novel location (i.e. the path changes). When a stick, rather than a hand, contacts the object, infants react only to the change in path. By one year, infants seem to understand even more complex facts about means–ends relations, relevant to Clause 3. For example, 12–14 month-olds recognize that actors understand means–ends relations and may take different alternative routes to obtain an end [43,44].

In terms of Clause 1, by 18 months, infants will 'read through' failed actions to infer the underlying intention of the actor [45]. When 18-month-olds see another person try and fail to pull apart an object for example, they will immediately pull apart the object themselves – something they will not do if they see a machine perform a similar action on the object. By two years, children explicitly and spontaneously explain goal-directed actions as the result of internally generated mental states, desires or intentions, that are designed to alter the world in particular ways [7].

Infants also generalize from their own interventions to those of others and vice-versa. For example, you can train three-month-old infants to reach for objects by giving them Velcro mittens that allow them to manipulate objects they would not otherwise be able to grasp [46]. Infants who received such training generalized from their own interventions and were more likely to understand the directed reaches of others. Conversely, the extensive literature on early imitation shows that nine-month-old

infants who see another person perform a novel intervention (i.e. an experimenter touching the top of a box with his head to make the box light up) will adopt that intervention themselves – the babies will put their own heads on the box [47].

Learning from combinations of conditional probabilities and interventions

We have seen that infants and young children seem to conceive of their own and others interventions in a distinctive way that might support causal learning. The crucial aspect of causal Bayes nets, however, is that intervention and conditional probability information can be coherently combined and inferences can go in both directions. Animals have at least some forms of the ability to infer conditional probabilities, and even conditional independencies, among events – as in the phenomenon of blocking in classical conditioning [48]. They also have at least some ability to infer causal relations between their interventions and the events that follow them, as in operant conditioning and trial and error learning. However, there is, at best, only very limited and fragile evidence of non-human animals' ability to combine these two types of learning in a genuinely causal way [49,50]. Why is it that when Pavlov's dogs associate the bell with food, they don't just spontaneously ring the bell when they are hungry? The animals seem able to associate the bell ringing with food, and if they are given an opportunity to act on the bell and that action leads to food, they can replicate that action. Moreover, there may be some transfer from operant to classical conditioning. However, the animals do not seem to go directly from learning novel conditional independencies to designing a correct novel intervention. Moreover, surprisingly primates show only a very limited and fragile ability to learn by directly imitating the interventions of others, an ability that is robustly present in one-year-old humans [50].

By contrast, very young children solve causal problems in a way that suggests just this coordination of observation and action. Preschool children, for instance, can use contingencies, including patterns of conditional independence, to design novel interventions to solve causal problems. Three-year-olds in the blicket detector experiments use information about conditional independence to produce appropriate interventions (such as taking a particular object off the detector to make it turn off) that they have never seen or produced before. [35–37].

Even more dramatically, four-year-olds used patterns of conditional dependence to craft new interventions that required them to cross domain boundaries, and overturn earlier knowledge [37]. For example, children were asked beforehand whether you could make a machine light up by flicking a switch or by saying 'Machine, please go'. All of the children said that flicking the switch would work but talking to the machine would not. Then the children saw that the effect was unconditionally dependent on saying 'Machine, please go', but was independent of the switch conditional on the spoken request. When children were then asked to make the machine stop 75% said 'Machine, please stop'.

Most crucially, however, four-year-olds can also combine patterns of conditional dependence and intervention to infer causal structure and do so in a way that recognizes the special character of intervention. This kind of inference is naturally done by Bayes nets and is not a feature of other accounts of causal reasoning such as associationist [51,52] or causal power [53] accounts. Children can use such combinations of information to identify causal direction (Does X cause Y or does Y cause X?) and even to infer the existence of unobserved variables. They can even do so when the relations between the events are probabilistic rather than deterministic [29].

For example, four-year-olds were shown a 'puppet machine' in which two stylized puppets moved simultaneously. They were told that some puppets *almost* always, but not always, made others go. In one condition they saw the experimenter intervene to move puppet X, and puppet Y also moved simultaneously on five of six trials. On one trial the experimenter moved X and Y did not move. In the other condition children simply observed the puppets move together simultaneously five times, but on one trial the experimenter intervened to move X and Y did not move. The children accurately concluded that X made Y move in the first case, whereas Y made X move in the second [29].

Conclusion

Although much more research is necessary (e.g. see Box 1), it seems that infants and young children can detect patterns of conditional probability, understand the

---

Box 1. Questions for future research

Questions about conditional probability
• Can children distinguish only conditional probabilities of 1 and <1 or can they make finer distinctions? Are judgments of conditional independence possible in infancy?
• How do children get from frequency information to judgments of conditional probability? How do they deal with the problem of small sample sizes?

Questions about intervention
• Do children treat only human actions as interventions or can they recognize 'natural experiments'?
• Do children understand that actions must fulfill the criteria of the Intervention Assumption to count as interventions? Do they discount 'bad' interventions?

Questions about causal structure
• Can children use patterns of evidence to discriminate more complex causal structures (e.g. causal chains versus common causes versus common effects)? Can they use them to determine parameterizations of a graph (e.g. the strength of causal links, and whether they are deterministic, generative, inhibitory or interactive)?
• Can children use patterns of evidence to determine unobserved as well as observed causal structure, to discover new variables, or split or merge existing variables?
• How do children integrate spatial and temporal information with information about conditional probability and intervention?

nature of their own and others interventions, and to at least some extent, integrate conditional probability and intervention information spontaneously and without reinforcement.

Each of these abilities, by itself, provides a powerful foundation for learning of several kinds, not just causal learning. Significantly, for example, in at least one experiment infants treated the units that emerged from statistical auditory regularities as English words, that is, as genuinely linguistic representations that could be combined with others in a rule-governed way [54]. Infants might similarly use conditional probabilities of visual stimuli to segregate scenes into object representations, which can then be combined in a rule-governed way [55]. Furthermore, understanding and imitating the interventions of others, not only in simple action imitation but in more complex cases such as taking on the goals of others, provides infants with powerful tools for learning social behavior [47].

Recent work on the causal Bayes net formalism, however, suggests that combining these two types of learning provides particularly powerful tools for learning causal structure, of the kind encoded in intuitive theories, and provides a formal account of how this might be done. Elements of such learning appear to be in place in infancy, and these elements are clearly used to learn causal relations by early childhood.

References
1 Gopnik, A. and Meltzoff, A.N. (1997) *Words, Thoughts and Theories*, MIT Press
2 Gelman, S.A. and Raman, L. (2002) Folk biology as a window onto cognitive development. *Hum. Dev.* 45, 61–68
3 Flavell, J.H. (1999) Cognitive development: children's knowledge about the mind. *Annu. Rev. Psychol.* 50, 21–45
4 Harris, P.L. *et al*. (1996) Children's use of counterfactual thinking in causal reasoning. *Cognition* 61, 233–259
5 Hickling, A.K. and Wellman, H.M. (2001) The emergence of children's causal explanations and theories: evidence from everyday conversation. *Dev. Psychol.* 37, 668–683
6 Sobel, D.M. (2004) Exploring the coherence of young children's explanatory abilities: evidence from generating counterfactuals. *Br. J. Dev. Psychol.* 22, 37–58
7 Wellman, H.M. *et al*. (1997) Young children's psychological, physical, and biological explanations. In *The Emergence Of Core Domains Of Thought: Children's Reasoning About Physical, Psychological, And Biological Phenomena. (New Directions for Child Development, No. 75)* (Wellman, H.M. and Inagaki, K., eds), pp. 7–25, Jossey-Bass/Pfeiffer
8 Slaughter, V. and Gopnik, A. (1996) Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Dev.* 67, 2967–2988
9 Slaughter, V. *et al*. (1999) Constructing a coherent theory: children's biological understanding of life and death. In *Children's Understanding Of Biology And Health* (Siegal, M. Peterson, C. *et al*., eds), pp. 71–96, Cambridge University Press
10 Slaughter, V. and Lyons, M. (2003) Learning about life and death in early childhood. *Cogn. Psychol.* 46, 1–30
11 Ross, N. *et al*. (2003) Cultural and experimental differences in the development of folkbiological induction. *Cogn. Dev.* 18, 25–47
12 Glymour, C. and Cooper, G. (1999) *Computation, Causation, and Discovery*, AAAI/MIT Press
13 Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann
14 Pearl, J. (2000) *Causality*, Oxford University Press
15 Spirtes, P. *et al*. (1993) *Causation, Prediction, and Search (Springer Lecture Notes in Statistic)*, Springer-Verlag
16 Hausman, D.M. and Woodward, J. (1999) Independence, invariance and the causal Markov condition. *Br. J. Philos. Sci.* 50, 521–583
17 Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press
18 Silva, R. *et al*. (2003) Learning measurement models for unobserved variables. In *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*, AAAI Press
19 Richardson, T. and Spirtes, P. (2003) Causal inference via ancestral graph models. In *Highly Structured Stochastic Systems* (Green, P. *et al*., eds), Oxford University Press
20 Ramsey, J. *et al*. (2002) Automated remote sensing with near-infra-red reflectance spectra: carbonate recognition. *Data Mining and Knowledge Discovery* 6, 277–293
21 Shipley, B. (2000) *Cause and Correlation in Biology*, Oxford University Press
22 Glymour, C. (2001) *The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology*, MIT Press
23 Glymour, C. (2003) Learning, prediction and causal Bayes nets. *Trends Cogn. Sci.* 7, 43–48
24 Glymour, C. and Cheng, P. (1999) Causal mechanism and probability: a normative approach. In *Rational Models of Cognition* (Oaksford, K. and Chater, N., eds), pp. 295–313, Oxford University Press
25 Rehder, B. and Hastie, R. (2001) Causal knowledge and categories: the effects of causal beliefs on categorization, induction, and similarity. *J. Exp. Psychol. Gen.* 130, 323–360
26 Steyvers, M. *et al*. (2003) Inferring causal networks from observations and interventions. *Cogn. Sci.* 27, 453–489
27 Waldmann, M.R. and Hagmayer, Y. (2001) Estimating causal strength: the role of structural knowledge and processing effort. *Cognition* 1, 27–58
28 Gopnik, A. and Glymour, C. (2002) Causal maps and Bayes nets: a cognitive and computational account of theory-formation. In *The Cognitive Basis of Science* (Carruthers, P. *et al*., eds), pp. 117–132, Cambridge University Press
29 Gopnik, A. *et al*. (2004) A theory of causal learning in children: causal maps and Bayes nets. *Psychol. Rev.* 111, 3–32
30 Aslin, R.N. *et al*. (1998) Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* 9, 321–324
31 Saffran, J.R. *et al*. (1996) Statistical learning by 8-month old infants. *Science* 274, 1926–1928
32 Saffran, J.R. *et al*. (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52
33 Fiser, J. and Aslin, R.N. (2002) Statistical learning of new visual feature combinations by infants. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15822–15826
34 Kirkham, N.Z. *et al*. (2002) Visual statistical learning in infancy: evidence of a domain general learning mechanism. *Cognition* 83, B35–B42
35 Gopnik, A. *et al*. (2001) Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Dev. Psychol.* 37, 620–629
36 Sobel, D.M. *et al*. (2004) Children's causal inferences from indirect evidence: backwards blocking and Bayesian reasoning in preschoolers. *Cogn. Sci.* 28, 3
37 Schulz, L. and Gopnik, A. (2004) Causal learning across domains. *Dev. Psychol.* 40, 162–176
38 Baldwin, D.A. and Baird, J.A. (2001) Discerning intentions in dynamic human action. *Trends Cogn. Sci.* 5, 171–178
39 Baldwin, D.A. *et al*. (1999) Infants parse dynamic action. *Child Dev.* 72, 708–717
40 Woodward, A.L. (1998) Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34
41 Woodward, A. and Sommerville, J.A. (2000) Twelve-month-old infants interpret action in context. *Psychol. Sci.* 11, 73–77
42 Phillips, A. *et al*. (2002) Infants' ability to connect gaze and emotional expression to intentional action. *Cognition* 85, 53–78

43 Gergely, G. *et al*. (2002) Rational imitation in preverbal infants. *Nature* 415, 755

44 Gergely, G. *et al*. (1995) Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193

45 Meltzoff, A.N. (1995) Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol.* 31, 838–850

46 Woodward, A.L. *et al*. (2001) How infants make sense of intentional action. In *Intentions and Intentionality: Foundations of Social Cognition* (Malle, B. *et al*., eds), pp. 149–169, MIT Press

47 Meltzoff, A.N. and Prinz, W., eds). (2002). Cambridge University Press

48 Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Theory and Research* (Black, A.H. and Prokasy, W.F., eds), pp. 64–99, Appleton-Century-Crofts

49 Povinelli, D. (2000) *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works*, Oxford University Press

50 Tomasello, M. and Call, J. (1997) *Primate Cognition*, Oxford University Press

51 Shanks, D.R. and Dickinson, A. (1987) Associative accounts of causality judgment. In *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 21) (Bower, G.H., ed), pp. 229–261, Academic Press

52 Shanks, D.R. (1985) Forward and backward blocking in human contingency judgement. *Q.J. Exp. Psychol. B* 37, 1–21

53 Cheng, P.W. (1997) From covariation to causation: a causal power theory. *Psychol. Rev.* 104, 367–405

54 Saffran, J. (2001) Words in a sea of sounds: the output of infant statistical learning. *Cognition* 81, 149–169

55 Fiser, J. and Aslin, R.N. (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* 12, 499–504