

Domain Generality and Specificity in Children's Causal Inference About Ambiguous Data

David M. Sobel
Brown University

Sarah E. Munro
University of California at Berkeley

In 5 experiments the authors examined children's understanding of causal mechanisms and their reasoning about base rates across domains of knowledge. Experiment 1 showed that 3-year-olds interpret objects activating a machine differently from a novel agent liking each object; children are more likely to treat the latter as indicating the objects with the causal property possessed an internal property. Experiment 2 suggested that 3-year-olds potentially use this mechanistic knowledge to reason about ambiguous data in terms of base rate information. Experiments 3, 4a, and 4b showed that these inferences are not the result of children being more interested in an agent's desires. Instead, children integrate domain-specific knowledge (i.e., reasoning about an agent vs. a machine) with the nature of that inference within that domain (i.e., reasoning about desires vs. other mental states). The authors suggest that a particular computational approach, based on Bayesian inference, best describes these inferences. This approach offers a description of how children might integrate domain-specific mechanism knowledge into a more general model of causal inference based on observing covariation data among events.

Keywords: causal reasoning, Bayesian inference, domain specificity, cognitive development

Numerous investigations have shown that young children can make causal inferences across a variety of domains, including physical knowledge (e.g., Bullock, Gelman, & Baillargeon, 1982; Spelke, Breinlinger, Macomber, & Jacobson, 1992), biological concepts (e.g., Inagaki & Hatano, 1993; Kalish, 1996), and folk psychology (e.g., Perner, 1991; Wellman, 1990). An open question in cognitive development is how children make these inferences.

One issue to consider when investigating how children make causal inferences is the relation between causality and *covariation*. Events or properties of objects that are causally related tend to co-occur, and causal relations can be inferred from covariation information (e.g., Cheng, 1997; Shanks, 1995). Children can make causal inferences on the basis of covariation patterns among physical events (Gopnik, Sobel, Schulz, & Glymour, 2001; Sobel & Kirkham, 2006), as well as across domains of knowledge (Schulz & Gopnik, 2004). As such, theories that focus primarily on this information have often been characterized as more domain general (e.g., Gopnik et al., 2004).

However, an equally important issue to consider is how objects and events affect each other—how children represent the *mechanisms* behind causal relations (e.g., Ahn, Kalish, Medin, & Gelman, 1995; Shultz, 1982). Infants appear sensitive to such knowledge: physical events that “touch” each other are registered differently than physical events that lack spatial contiguity (e.g., Leslie & Keeble, 1987; Oakes & Cohen, 1990; following work by Michotte, 1962). Infants also register differences between physical and psychological efficacy on the basis of the nature of the actor and the actions performed on an object (e.g., Johnson, Slaughter, & Carey, 1998; Meltzoff, 1995; Saxe, Tenenbaum, & Carey, 2005) or on the basis of patterns of motion among objects (e.g., Gergely, Nádasdy, Csibra, & Bíró, 1995; Kuhlmeier, Wynn, & Bloom, 2003). Because children develop mechanistic knowledge across domains along different trajectories, theories that focus on mechanism information are often considered more domain specific.

Various accounts of causal reasoning have proposed ways of integrating domain-general principles of covariation with domain-specific mechanistic knowledge (e.g., Bullock et al., 1982; Carey & Spelke, 1994; Gopnik & Meltzoff, 1997). The majority of these accounts, however, lack a formal description of how mechanistic knowledge is integrated with data. Following proposals in cognitive psychology (Griffiths & Tenenbaum, 2005; Tenenbaum & Griffiths, 2003; Tenenbaum, Griffiths, & Niyogi, 2007), we describe this integration by appealing to a model of Bayesian inference that incorporates the child's existing causal knowledge. This account is meant as a computational-level description (cf. Marr, 1982) of children's causal inferences and should not be taken as the process by which children consciously engage in such inferences.

In this article, we first describe the Bayesian framework as well as an experiment on children's causal inferences designed to test this approach (Sobel, Tenenbaum, & Gopnik, 2004). This experiment found a difference between 3-year-olds' and 4-year-olds'

David M. Sobel, Department of Cognitive and Linguistic Sciences, Brown University, and Sarah E. Munro, Helen Wills Neuroscience Institute, University of California at Berkeley.

This work was supported by National Science Foundation Grant DLS-0518161 awarded to David M. Sobel. We would like to thank all of the parents and children who participated in this research. We would also like to thank Esra Asku, Emily Blumenthal, Lorraine Fryer, Emily Hopkins, Philip Parker, Emily Stoddard, Kristen Sylvester, and Lea Ventura, who helped with data collection and coding, and Tom Griffiths, Natasha Kirkham, Josh Tenenbaum, and Amanda Woodward, who provided insightful feedback on this article.

Correspondence concerning this article should be addressed to David M. Sobel, Department of Cognitive and Linguistic Sciences, Box 1978, Brown University, Providence, RI 02912. E-mail: Dave_Sobel@brown.edu

inferences. Our hypothesis is that this development can be explained by positing that 4-year-olds but not 3-year-olds possess the mechanism knowledge necessary to make inferences consistent with the Bayesian description. The five experiments presented here were conducted to examine how changing the nature of the mechanism by which children interpret covariation data allows 3-year-olds to reason consistently with the Bayesian description.

Bayesian Inference, Mechanism Knowledge, and Young Children¹

A Bayesian learner begins with a set of hypotheses (h). Each hypothesis (h_1 , h_2 , and so on) is assigned a prior probability, $p(h)$. These hypotheses and priors reflect the learner's existing mechanism knowledge about what could explain potential data and how events are related. After observing data (d), the learner computes the posterior probability that each hypothesis is the actual causal structure of the data he or she observed, $p(h|d)$. The learner does this using Bayes' rule, shown in Equation 1. In this equation, $p(h)$ is the prior, and the likelihood value $p(d|h)$ describes how probable the data are if the hypothesis in question is the actual causal structure of the world.

$$p(h|d) = \frac{p(d|h)p(h)}{\sum_{h' \in H} p(d|h')p(h')} \quad (1)$$

Sobel et al. (2004) used an empirical procedure that illustrates how a causal inference task can be formulated as a problem of Bayesian inference. They presented 3- and 4-year-olds with a machine that lit up and played music when particular objects were placed upon it. Children were shown a set of identical blocks and were told that "blickets make the machine go." They then observed that blickets were either *rare* or *common*—that either 2 or 10 of the first 12 blocks taken from the set activated the machine (and received the appropriate *blicket* or *not-blicket* label from the experimenter). Then, two new blocks (A and B) were taken from the set and placed on the machine together, which activated. Children were then shown that Block A activated the machine by itself. They were asked whether each block was a blicket. All children categorized Block A as a blicket, but the status of Block B is ambiguous (for reasons that will be described in detail later). Four-year-olds resolved this ambiguity using the base rate information they were shown. When blickets were rare, they categorized Object B as not a blicket ~75% of the time. When blickets were common, they categorized Object B as a blicket ~80% of the time. Three-year-olds, however, were insensitive to base rates: They categorized Object B as a blicket equally and often (~80% of the time) across the two conditions.

To apply Bayesian inference to this procedure, consider that there are four hypotheses potentially consistent with these data at the outset:

- h_1 : that neither object has the power to activate the machine (i.e., is a blicket)
- h_2 : that only A is a blicket
- h_3 : that only B is a blicket
- h_4 : that both A and B are blickets

These hypotheses result from particular pieces of knowledge that 3- and 4-year-olds clearly possess. This knowledge includes

(a) categorization—an understanding that there are blocks and blicket machines, which are different kinds of objects; (b) knowledge of temporal priority—placing a block on the machine causes the machine to activate, not the reverse; and (c) an understanding of spatial independence—blocks (and hence, blickets) exist independent in space from one another.

Given this hypothesis space, a rational way of assigning priors is by observing the base rate of objects with causal efficacy. Thus, if there are few blickets in the world, then hypotheses that propose fewer blickets will have a higher prior and those that propose there are many blickets will have a lower prior. If there are many blickets in the world, the reverse is true.

The data d that children observe is that Objects A and B together activate the machine and then that Object A activates the machine by itself. Given these data, a learner must determine the likelihood functions for each hypothesis, $p(d|h)$. Since the data indicate that A is a blicket, hypotheses h_1 (neither A nor B is a blicket), and h_3 (only B is a blicket) are inconsistent with the data, and thus their likelihoods $p(d|h) = 0$. Calculating these likelihoods of h_2 and h_4 relies on another piece of mechanism knowledge—how blickets are related to the machine. The ambiguity in these data comes from recognizing what Tenenbaum and Griffith (2003) called an *activation law*: the machine's activation is produced by a stable mechanism inherent to blickets that always makes the machine go. This piece of knowledge posits that both h_2 and h_4 are consistent with the data, and their $p(d|h) = 1$. Block B could have the power to activate the machine on its own, or the machine activated when B was on it because A was on it as well.

To see why the ambiguity is only present when blickets always activate the machine, consider if blickets do not always activate the machine. In this case, even though Block A is clearly a blicket, it might have failed to produce its efficacy on the trial in which Blocks A and B activated the machine together. This means that the likelihood that Block B might be a blicket is greater than just the base rate of blickets existing in the world. Of course, children need not make such a sophisticated inference—it is likely that if they do not recognize that blickets are connected to machine's activation by some mechanism, they might respond on the basis of just the positive association between Object B and activation in these data. This would also suggest that B is a blicket.

If these likelihood values are inserted into the formula shown in Equation 1, the posterior probabilities, $p(h|d)$, of h_1 and h_3 both equal 0, while those of h_2 and h_4 are proportional to the prior, $p(h)$ —if there are few blickets in the world, h_2 will have a high posterior (making Block B unlikely to be a blicket); if there are many blickets, h_4 will have a high posterior (making Block B likely to be one).

While this description potentially explains 4-year-olds' inferences, 3-year-olds clearly did not reason in this manner. We hypothesized that children develop an understanding of the acti-

¹ This section was developed in collaboration with Tom Griffiths and Josh Tenenbaum, and we are greatly indebted to them for developing this model. This section is intended as a brief introduction to this model and to Bayesian inference in general. We will not describe this model in all its details. An initial attempt at describing these details was done by Tenenbaum & Griffiths (2003). A more rigorous description was presented in Griffiths and Tenenbaum (2007) and Griffiths et al. (2009).

vation law that connects blickets to the machine when they are between the ages of 3 and 4, which potentially explains this developmental difference. If 3-year-olds recognize that an object's causal efficacy is the result of a stable mechanism (an activation law for that causal relation), we hypothesized that their inferences would be more consistent with the Bayesian description.

Mr. Blicket and the Present Experiments

To test these hypotheses, we conducted Experiment 1 to consider whether 3-year-olds recognized different activation laws across different domains of knowledge. Specifically, we examined how children reasoned about the relation between an object's causal efficacy and its internal properties when the causal relation was presented as an object activating a machine or as an agent liking an object.

Why examine insides? Many causal mechanisms have obvious perceptual correlates: a part or feature that allows the child to see how the object produces its efficacy. Therefore, it is difficult to discern whether children make inferences on the basis of the perceptual information they have available to them or on the basis of a more conceptual understanding about mechanisms. We wanted to examine whether 3-year-olds grasped that a mechanism relates an object with its causal efficacy by considering how children reasoned about a *nonobvious* property of that object and specifically when that property is contrasted against the object's external perceptual appearance. Because 3-year-olds are notoriously bad at generating mechanistic explanations (e.g., Bullock et al., 1982), we used a manipulation that involved a more implicit expression of mechanism. In particular, Sobel, Yoachim, Gopnik, Meltzoff, and Blumenthal (2007) found that 4-year-olds, but not 3-year-olds, recognized that an object's causal efficacy (about whether it would activate a machine) was related to its insides, even in cases of perceptual contrast. Four-year-olds also recognized that if an object's insides were transferred to another object, the object's causal properties would transfer as well. This suggests that 4-year-olds understood that insides were responsible for an object's causal efficacy, potentially indicating their grasp of an activation law that allows them to assign likelihood functions in the manner described above. Such development reflects 3- and 4-year-olds' developing knowledge of several other kinds of non-obvious mechanisms behind physical causal relations (Bullock et al., 1982; Shultz, 1982).

Why examine an agent's desires? Various studies focusing on children's understanding of causal mechanisms have emphasized that infants treat physical and psychological causal relations differently (e.g., Meltzoff, 1995; Saxe et al., 2005). This suggests that we might want to contrast how children reason about machines with a social-cognitive mechanism. Three-year-olds recognize that another's desires can be different from their own and that others' preferences can relate to nonobvious properties of objects (e.g., Fawcett & Markson, 2007; Gopnik & Slaughter, 1991; Repacholi & Gopnik, 1997). Three-year-olds also know that if someone wants something, they will continue to act to fulfill that particular desire (Wellman & Woolley, 1990).

One's desires can certainly be based on external visual appearance (i.e., an individual can like a blue sweater and not a yellow one because of color), but desires can also be based on nonobvious properties (i.e., an individual might like one blue sweater over

another because of a nonobvious preference—for example, one has to be dry cleaned and the other is machine-washable or one stays soft with use while the other becomes itchy). Fawcett and Markson (2007) in particular suggested that 3-year-olds might be sensitive to the latter information. If one confederate liked the same toys as a child and another confederate had different preferences, children were more likely to play with a novel object preferred by the first confederate over one preferred by the second. Children registered that there might be something about that particular novel object that would be preferable to them on the basis of appreciating another's desires. From this, we hypothesized that 3-year-olds might generally understand that there is a mechanism—a reason—behind an agent's desires. We did not believe that there is something special about desire as a mental state that promotes causal inferences consistent with the Bayesian framework. Rather, we believed that 3-year-olds might be more likely to understand that desires have a stable mechanism—an activation law—in this context.

In Experiment 1, children were assigned either to a *machine* condition or *agency* condition. Children in the machine condition were shown that some objects activated a machine. In the agency condition, children were introduced to *Mr. Blicket*—a blicket machine with a set of cardboard eyes that “talked” to the experimenter and the child by activating spontaneously (following a manipulation used by Johnson et al., 1998). These children were told that when Mr. Blicket activated in response to an object being placed on him, he liked the object. This manipulation allowed us to keep the data children observed practically identical between the two conditions, while changing how the objects' causal properties were interpreted. Three-year-olds were expected to recognize that Mr. Blicket's preference was caused by a stable property inherent to the object, making them more likely to relate objects' causal efficacy with their insides in the agency condition.

Experiment 2 examined whether this agency manipulation would allow 3-year-olds to use base rate information to resolve ambiguous data. Following the procedure used by Sobel et al. (2004), we showed 3-year-olds that Mr. Blicket liked many or few objects from an identical set and then we showed them data that would be ambiguous if children recognized there was a stable mechanism between object and efficacy (specifically, that he liked two new Objects A and B and then that he liked just Object A). If 3-year-olds infer that those desires result from a property of the objects, they might interpret this property as a mechanism linking the object with the agent's desire. This would allow them to assign the appropriate likelihood functions and respond with the base rate in mind. If they do not, they might just use the associative information inherent in the procedure.

Experiment 3 replicated the procedure in Experiment 2 with a spontaneously activating machine to match the unfamiliarity and saliency of the agency condition. In Experiments 4a and 4b, we then examined other mental states that are not understood by 3-year-olds as well. In these experiments, we posited that most 3-year-olds would not understand the mechanism underlying the causal relations they observe. Children were presented with a machine or an agent who was thinking or who was knowledgeable about objects—mental states 3-year-olds often fail to grasp. We expected inferences to be different from those in Experiment 2 and not reflect the base rate information they observed. Rather, we thought that children might be more likely to respond on the basis of associative information inherent in the data—not that the data were

ambiguous or that the base rate information was relevant. This would suggest that it is not reasoning about agency that promotes inferences consistent with the Bayesian description but rather an understanding of the mechanism underlying the causal relations that were observed. In Experiment 2 and these last three experiments, after considering whether children used base rate information, we gave them a measure that considered their understanding of the mechanism (by administering a procedure similar to that used in Experiment 1). We believed that the way that they construed the causal relations they observed would affect their inferences, so we predicted that children who used base rate information would be more likely to appreciate that an object's causal efficacy is related to a stable internal property.

Experiment 1

Experiment 1 examined under what conditions 3-year-olds understood that objects with shared causal efficacy shared a nonobvious property. Half of the children were given an *agency* procedure, in which they were introduced to Mr. Blicket, a box with cardboard eyes that activated contingent on human interaction. The other children were given the *machine* condition, in which they saw the same box without eyes activate spontaneously but not contingent on the experimenter or child. In the agency condition, children were told that Mr. Blicket liked an object when he activated, while in the machine condition, children were told that the “machine was going” when it activated.

Children were shown evidence that the perceptual features of objects conflicted with their causal properties. Given these data, children were asked which of these factors—external perceptual features of causal properties predicted a nonobvious internal property of the objects. Our question was whether 3-year-olds would treat the two causal properties differently. We speculated that 3-year-olds would appreciate the relation between causality and insides in the agency condition more than the machine condition.

Method

Participants

Thirty-two 3-year-olds (20 girls, 12 boys; age $M = 41.13$ months, $SD = 3.08$) were recruited from flyers posted at urban-area preschools. Eight other children were tested but not included in the final sample: 2 children were nonnative speakers of English, 1 child refused to participate, and 5 children's data could not be used because of experimental error. The ethnic breakdown of the sample was as follows: 28 children were White, 3 children were Hispanic, and 1 was Middle Eastern. No information about parental education or socioeconomic status (SES) was collected.

Materials

The machine was a 20.3 cm \times 15.2 cm \times 7.6 cm box, made of black plastic with a white top, which was wired to a set of LEDs (see Figure 1a). The machine could activate via a remote control hidden from the child, and it could be enabled such that any object that touched its top would cause it to activate as soon as the object came into contact. Activation would cease when that object was removed. When it activated, the machine's top panel illuminated (by a set of green LEDs) and played music. The machine was battery-powered, so no external cords ran to or from it. This design

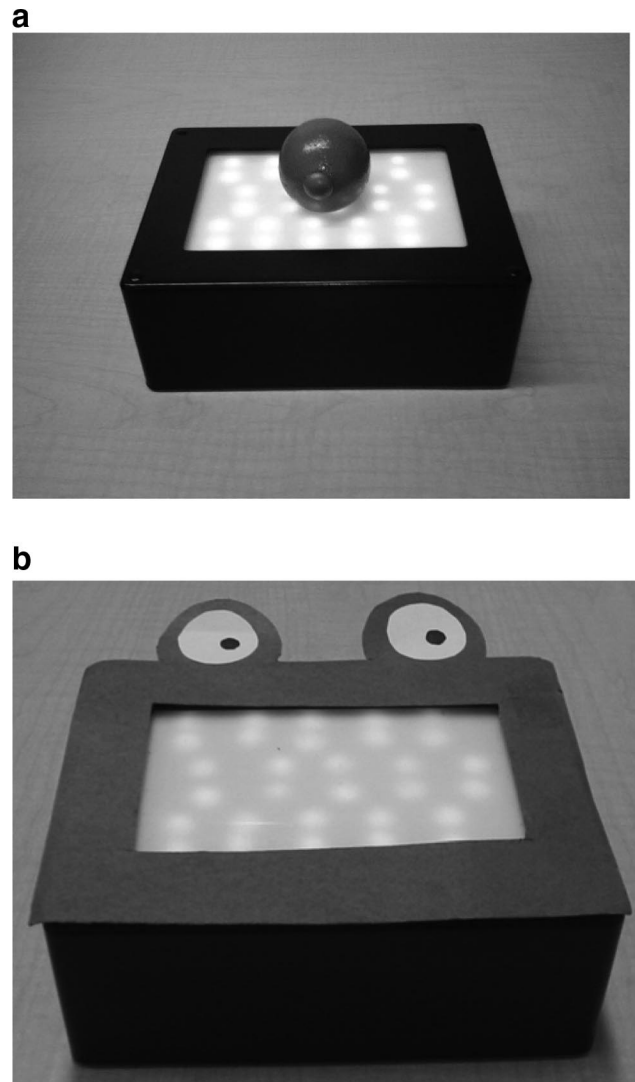


Figure 1. (a) The Blicket detector (in an active state) used in Experiment 1 (machine condition) and Experiment 3. (b) Mr. Blicket used in Experiment 1 (agency condition) and Experiments 2, 4a, and 4b. The block on the machine in Figure 1a was one of the stimuli used throughout the experiments.

facilitated the illusion that objects activated the machine. The box was presented as either a machine or as “Mr. Blicket.” When presented as Mr. Blicket, the top was covered by blue cardboard that resembled eyes, cut such that the white top of the box appeared to be Mr. Blicket's mouth. Figure 1b shows the detector as Mr. Blicket in an active state.

Four sets of 3 wooden blocks of various shapes, colors, and sizes were used. None of these 12 blocks was red or green. A 1.3 cm (diameter) \times 2.5 cm (depth) hole was drilled into each block. Eight of the blocks contained a large red map pin inside the hole; the remaining 4 were empty. Each block had a removable dowel that covered the hole, making it impossible to see whether it contained a map pin without removing the dowel. These dowels were only removed by the experimenter during the course of the

procedure. The map pins were placed deep enough so that the child could see them if the dowel was removed but could not remove them from the block. In each set, 2 blocks had identical external perceptual features. One of these contained a map pin (the target); the other did not (the externally similar object). The third block differed from the other 2 in both color and shape but did contain a map pin (although children never observed this). Figure 2 showed an example set, without the dowels covering the cavities. Four other blocks of different sizes, shapes, and colors were used in the warm-up.

Procedure

Children were brought into a room with an experimenter with whom they were familiar. They were given a pretest establishing that they would point to blocks of particular sizes and colors (this was done to establish that they would not touch the test objects; pilot work found this prevented children from spontaneously removing the dowels during the test phase). Children were randomly assigned to one of the two conditions. In the agency condition, the experimenter introduced the box as Mr. Blicket and then conducted a dialogue with it. Mr. Blicket responded to questions and comments by activating spontaneously. The experimenter also encouraged the child to interact with Mr. Blicket by saying hello, in which case Mr. Blicket activated in response. The children were then told that they were going to play a game in which Mr. Blicket would tell them whether he liked an object.

In the machine condition, the box was introduced to the child as a Blicket machine. The experimenter spoke only to the child, first by making small talk and then by telling the child that they were going to play a game to see which blocks made the machine go. During this time, the detector activated spontaneously as in the agency condition. No explanation for activation was provided, and activation was not contingent on the experimenter's speech.

In both conditions, the trials were similar. In each of the four trials, a set of two identical blocks and one unique block was placed in front of the machine/Mr. Blicket in a random configuration. Each block from the experimenter's right to left was individually placed on the machine/Mr. Blicket, which activated for one of the identical blocks and the unique block. In the agency condition, the experimenter said, "Look, Mr. Blicket likes this one" when the box activated or "Look, Mr. Blicket does not like this one" when it did not. In the machine condition, the child was

told either "Look, the machine is going" or "Look, the machine is not going." Each block was demonstrated twice, with the same outcome. The box did not spontaneously activate during this demonstration.

Children were then told that the objects all had "doors" (referring to the dowels) on them. The member of the identical pair that had activated the machine was picked up, and its dowel was removed, revealing that it contained a red map pin. The experimenter then remarked, "Look, this one is made of red stuff" and asked the child to point to "another one made of red stuff." After the child responded, the objects were removed from the table, and the next trial began. Children received the four trials in one of four quasi-random orders, counterbalanced across children.

Results

For each trial, children were given a score of 1 if they pointed to the block with the same causal efficacy as the target and a score of 0 if they responded on the basis of perceptual similarity. Children's scores did not differ across the four trials, Cochran's $Q(3, N = 32) = 0.69, ns$. These scores were summed. A preliminary analysis also revealed no effect of trial order on this summation score, Kruskal-Wallis test, $\chi^2(3, N = 32) = 2.08, ns$. The distribution and summary data of these scores are shown in Table 1.

Children chose the perceptually distinct object on 2.81 out of 4 (70%) of the trials in the agency condition, significantly more than in the machine condition (in which they made this response on 1.63 out of 4, or 41% of the trials), $t(30) = 3.24, p = .003$, Cohen's $d = 1.15$. We also categorized response patterns into three groups. Children were categorized as making *causal* responses if they chose the externally distinct object on three or four trials. Children were categorized as making *perceptual* responses if they chose the externally distinct object on none or one of the trials (and hence, made three or more perceptual responses). Children were categorized as being *neutral* if they made exactly two perceptual and two causal responses. The distribution of responses differed between the two groups, $\chi^2(2, N = 32) = 6.54, p = .03$, Cramer's $\phi = 0.45$. Specifically, 11 of 16 children (69%) in the agency condition made causal responses compared with only 4 children (25%) in the machine condition. The percentage of children who made causal responses in the agency condition was significantly greater than percentage that would be expected by chance (31.25%), binomial

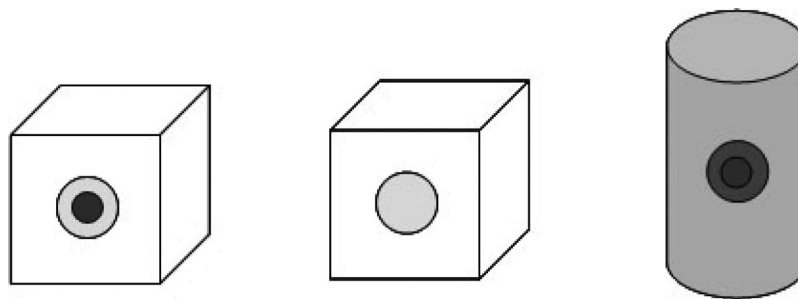


Figure 2. Example of a stimulus set used in Experiment 1 and in the internal properties measures in Experiments 2, 3, 4a, and 4b. In Experiment 1, four of these sets were constructed. In Experiments 2, 3, 4a, and 4b, we used three of those four sets. The figure shows these blocks with their dowels off, but children usually observed these blocks with their internal properties hidden behind the dowels.

Table 1
Distribution and Mean of Causal Responses in Experiment 1

Condition	No. of causal responses					Statistics	
	0	1	2	3	4	<i>M</i>	<i>SD</i>
Agency	0	2	3	7	4	2.81	0.98
Machine	3	4	5	4	0	1.63	1.09

Note. Means out of a possible 4.00.

test, $p < .001$. This was not the case in the machine condition, binomial test, *ns*.

Discussion

Three-year-olds were more likely to relate a nonobvious, internal property of an object with its causal efficacy when they believed the causal relation they observed was an agent's preference as opposed to the workings of a physical machine. Even though children observed the same data across the conditions, they appeared to construe those data differently and were less likely to use perceptual similarity in the agency condition.

These data do not confirm that 3-year-olds believed that the objects' insides were responsible for their causal efficacy. It is possible that children believed that the object's material was the mechanism underlying Mr. Blicket's preference. It is also possible that children simply recognized that the objects' external perceptual features were not the reason behind these preferences. We set up the agency condition such that objects were placed on what should have been considered Mr. Blicket's "mouth," which might indicate that objects have a particular taste, but the exact nature of how children interpret the relation between Mr. Blicket and the objects is an open question.² A conservative conclusion from Experiment 1 is that 3-year-olds interpret the same covariation information between the agency and machine conditions differently, potentially on the basis of the difference in mechanism information inherent in the manipulation.

Experiment 2

Given that Experiment 1 showed that 3-year-olds were more likely to relate an object's causal properties with a potential mechanism in the agency condition, in Experiment 2, we examined whether 3-year-olds would make causal inferences more consistent with the Bayesian description discussed in the introduction when this agency manipulation was used. We introduced 3-year-olds to Mr. Blicket and showed them that he liked either many or few objects from a set. The children then saw two new objects activate Mr. Blicket and later saw one of those objects alone activate him. If 3-year-olds have an activation law for desire—if they recognize that objects cause Mr. Blicket to say he likes them because of a stable mechanism that produces that efficacy—then they might treat these data as ambiguous and resolve the ambiguity by appealing to the base rate of the objects that Mr. Blicket likes. We hypothesized that 3-year-olds would respond in this manner, more like the 4-year-olds and less like the 3-year-olds in Sobel et al.'s (2004) investigation, on which this procedure was based.

Further, we predicted that an individual child's knowledge of the relation between objects' causal efficacy and their nonobvious properties would be related to their use of the base rate information. Children were also given a procedure similar to that used in Experiment 1 to see whether each child linked Mr. Blicket's desires with a nonobvious property (i.e., the insides) of the blocks. Children whose responses were consistent with the base rate should be more likely to recognize the relation between an object's insides and its causal properties (i.e., knowledge that might indicate a potential activation law). This is particularly important in the rare condition, since a response consistent with the base rate (i.e., saying Block B is not efficacious) is inconsistent with the associative information that children observe.

Method

Participants

Forty 3-year-olds (20 girls, 20 boys; age $M = 43.33$ months, $SD = 3.08$) were recruited in the same manner as Experiment 1. Eleven children were also tested, but their data were not included in the final analysis. Six failed the control procedure (discussed in a later section): 2 children refused to participate, 2 children were nonnative speakers of English, and 1 child had data that was not usable because of experimental error. The ethnic breakdown of the sample was as follows: 34 children were White, 1 child was Asian, 1 was African-American, and 4 were of mixed descent. No information about parental education or SES was collected.

Materials

The same Mr. Blicket and three of the four block sets from Experiment 1 were used. Eighteen identical 3.5 cm \times 7.7 cm blue cylindrical blocks and three cardboard boxes were also used. One box, which held the 18 blocks at the start of the procedure, was 15.8 cm \times 15.8 cm \times 15.3 cm. The other two boxes were 13.5 cm \times 8.8 cm \times 27.2 cm. One had the words "Mr. Blicket likes it" written on it; the other had the words "Mr. Blicket does not like it" on it.

Procedure

All children were given a base rate measure and an internal property measure. Children also received a set of unrelated measures and participated in another experiment in between these two measures. Children always received the base rate measure first, since presenting it after the internal property measure would have affected the base rate of objects Mr. Blicket liked.

Base rate measure. The procedure was a modification of that used by Sobel et al. (2004). Children were introduced to Mr. Blicket in the same manner as in the agency condition in Experiment 1. They were told that the experimenter had a box of toys (the box of 18 identical blocks). The experimenter asked the child

² There is good reason to believe that children might not need to be able to articulate such a mechanism in order to reason appropriately about it. There are several cases in which children (and adults) reason appropriately about mechanism information, yet are completely unable to articulate such information (see e.g., Rosenbitz & Keil, 2002).

whether he or she saw the toys and then turned them to Mr. Blicket and asked whether he saw the toys, at which point Mr. Blicket activated spontaneously, which the experimenter interpreted as a “yes” answer. The experimenter then said, “OK, Mr. Blicket. I want you to tell us which ones you like. Can you do that?” Again, Mr. Blicket spontaneously activated. The experimenter then said, “Great”; removed two blocks from the box; said to the child, “OK, let’s try this one”; and placed one block on Mr. Blicket. When this block was placed on him, Mr. Blicket activated, and the experimenter said, “Look, he likes this one.” The experimenter then brought out the small box with the words “Mr. Blicket likes it” on it and said, “I have a box. Let’s put all of the things that he likes into this box.” The experimenter read the label to the child and performed this action. The experimenter then picked up the other block, said, “OK, let’s try this one,” and placed it on Mr. Blicket. Mr. Blicket did not activate, and the experimenter said, “Look, he does not like this one. I have another box (the box with the words “Mr. Blicket does not like it”). Let’s put all of the things that he does not like into this box,” and the experimenter read the label on the box and performed this action.

The experimenter then pointed to the two boxes and asked the children whether they remembered the purpose of each box (i.e., whether one was for things Mr. Blicket liked and one was for things he did not like). Children received corrective feedback if they answered incorrectly. Then, the experimenter took the two blocks that had already been categorized; said, “Let’s see what happens when we put these on Mr. Blicket together”; and placed them on Mr. Blicket together. Mr. Blicket activated, and the experimenter said, “Look at that; he likes them together.” The experimenter then said, “But that’s because he liked this one” (leaving only the object that Mr. Blicket had previously showed that he liked on Mr. Blicket, who continued to activate). This object was removed, and the other object was placed on Mr. Blicket, who did not respond. The experimenter said, “And this one, he does not like.” Then these two objects were placed back in their respective boxes. This demonstration ensured that children would recognize that Mr. Blicket activated if at least one object that he liked was present.

Children were then shown the effects of 10 training blocks (presented in pairs). Approximately half of the children ($n = 25$) were assigned to the *rare* condition and were shown that 1 of these 10 blocks activated Mr. Blicket (for an overall base rate of 2 of 12, or 1/6). The remaining children ($n = 15$) were assigned to the *common* condition and were shown that 9 of these 10 blocks activated Mr. Blicket (for an overall base rate of 10 of 12, or 5/6). After seeing the efficacy of each block, children were asked to place it into the appropriate box. Corrective feedback was provided if necessary.

After this demonstration, the experimenter said, “Look at what we did. We took a whole bunch of blocks out of here (the original box) and asked Mr. Blicket if he liked each one.” In the rare condition, the experimenter said, “He liked a few of them (pointing to the box of blocks that he liked, which contained 2 blocks). But, most of them, he did not like (pointing to the box of blocks that he did not like, which contained 10 blocks).” In the common condition, this was reversed to remind children about the base rate of objects Mr. Blicket liked.

Children were then given three trials: the *test*, *baseline*, and *control* trials. In the test trial, 2 new blocks were brought out of the

original box (A and B) and placed on Mr. Blicket together, who activated. Then Block A alone was placed on Mr. Blicket, who again activated. Children were asked to categorize these objects. Our first question was whether children in the two conditions treated Block B differently.

In the baseline trial, 2 new blocks were brought out (X and Y). Both were placed on Mr. Blicket together, who activated. Children were then asked whether Mr. Blicket liked each of these blocks. Would children treat these objects differently from Block B? The compound presentation of X and Y together should raise the likelihood that each has efficacy, since he must like at least one of them. But if children understand that Block B’s activation could have been explained by Block A’s efficacy, then at least in the rare condition they should believe he is less likely to like Block B than Blocks X and Y.

Finally, in the control trial, two new objects were placed on Mr. Blicket, who activated for one but not the other. Children were asked to put these 2 blocks into the boxes. This trial was used to ensure that children understood the task. If children made an error on the control trial, they were not included in the final sample. Six children were excluded for this reason (3 children in each condition). The spatial location of the 2 blocks in each trial was counterbalanced.

Internal property measure. After participation in an unrelated experiment and a break, children were introduced again to Mr. Blicket, and the experimenter had a brief conversation with the child to remind him or her about Mr. Blicket. The procedure here paralleled the procedure used in the agency condition in Experiment 1. Children were told that the experimenter had some other toys and that they would see whether Mr. Blicket liked them. The same sets of toys from Experiment 1 were used, and the same procedure was followed, except that only three trials were presented (order randomized). Thus, children were asked whether objects’ shared causal efficacy was related to their sharing insides in light of a conflict with external perceptual features. Our goal was to consider whether the greater the frequency with which children made this claim could predict whether they used the base rate information in the previous measure.

Results

On the base rate measure, children required corrective feedback on an average of 0.43 out of a possible 10 training trials. This suggested that children understood the basic nature of the measure before being given the test trial. All children responded that Mr. Blicket liked Object A across both conditions on the test trial. Responses to Block B are shown in Table 2. In the rare condition, 11 of 25 children (44%) claimed that Mr. Blicket liked B. In the common condition, 14 of 15 children (93%) claimed that Mr. Blicket liked B, a significant difference, $\chi^2(1, N = 40) = 9.74, p < .005, \psi = 0.49$.

Children did not categorize Objects X and Y differently on the baseline trial, so these data were averaged together. In the rare condition, children claimed that Mr. Blicket liked these objects 80% of the time, significantly greater than Object B in the test trial, Wilcoxon signed rank test, $z = -3.29, p = .001, r = -0.47$. In the common condition, children claimed that Mr. Blicket liked these objects 100% of the time, no different from their treatment of Object B.

Table 2
Distribution of Responses on the Base Rate Measure Across Experiments

Experiment	Children (%) stating each object has efficacy	
	Object A (unambiguous)	Object B (ambiguous)
Experiment 2		
Rare condition ^a	100	44
Common condition ^b	100	94
Experiment 3 ^a	100	72
Experiment 4a ^a	100	72
Experiment 4b ^b	94	80

Note. In Experiment 2, children were told about the agent’s desires, which were either rare or common. In Experiment 3, children reasoned about a machine’s activation; in Experiment 4a, about an agent’s thoughts about the objects; and, in Experiment 4b, about whether the agent said that each object was a blicket. In all experiments, children observed that Objects A and B had efficacy together and then that Object A had efficacy alone. Children categorized each object as either having or not having causal efficacy.

^a *N* = 25. ^b *N* = 15.

Performance on the base rate measure was compared with performance on the similar procedure used by Sobel et al. (2004). There, sixteen 3-year-olds were introduced to a “blicket machine” and were told that “blickets make the machine go.” Children were told that blickets were either rare or common, were then given the same three test trials, and for each were asked whether each object was a blicket. In the test trial, all children claimed that Object A was a blicket across both conditions. In the common condition, most of the 3-year-olds (14 of the 16, or 88%) also claimed Object B was a blicket, no different from the 93% of children who did so in this experiment. In the rare condition, however, most of the 3-year-olds (13 of the 16, or 81%) also claimed that Object B was a blicket. This represents a significantly greater percentage than the 44% of the 3-year-olds who did so here, $\chi^2(1, N = 41) = 5.58, p = .018, \psi = 0.37$. This suggests that when 3-year-olds were presented with the same kind of causal inference about a physical device, they failed to use base rate information. We will expand on this finding in Experiment 3.

The internal properties measure was scored in the same manner as in Experiment 1. The children’s performance on the three trials did not differ, Cochran’s $Q(2, N = 40) = 0.38, ns$, so these data were summed together and are shown in Table 3. Overall, children made causal responses on 41% of the trials (an average score of 1.23 out of a possible 3, $SD = 1.07$), and level of performance was not different between the rare and common conditions, Mann–Whitney $U = 189.00, z = -.015, ns$. The overall pattern of response was also not different from chance (50%, since the child chose between only the perceptually identical or causally similar options), $t(39) = -1.62, ns$. This is different from Experiment 1, in which responses in the agency condition were significantly greater than chance levels.

When these data were considered in light of performance on the base rate measure, a different response pattern emerged. Across the rare and common conditions, 28 of the 40 children responded in a manner consistent with the base rate: They stated that Mr.

Blicket did not like Object B in the rare condition and did like Object B in the common condition. These 28 children made causal responses on 51% of the internal property trials (an average score of 1.54 out of a possible 3, $SD = 1.10$). The remaining 12 children, whose responses were inconsistent with the base rate, made causal responses on 17% of the internal property trials (mean score of 0.50 out of a possible 3, $SD = 0.52$). These scores were significantly different from each other, Mann–Whitney $U = 78.00, z = -2.76, p < .01, r = -0.44$. However, the scores of the 28 children whose responses were consistent with the base rate were not different from chance.

We were also interested in responses in the rare condition alone, since children in the common condition might have stated that Mr. Blicket liked Object B (the response consistent with the base rate) not because they detected the ambiguity in the data but just on the basis of positive association. Fourteen of the 25 children in the rare condition claimed that Mr. Blicket did not like Object B (consistent with their recognizing the ambiguity of the data and using the base rate information). These children made causal responses on 57% of the internal property trials (an average score of 1.71 out of a possible 3, $SD = 0.91$). The 11 children who claimed that Mr. Blicket liked Object B (inconsistent with recognizing the ambiguity of the data and using the base rate information) only made causal responses on 18% of the trials (an average score of 0.55 out of a possible 3, $SD = 0.52$). This was a significant difference, Mann–Whitney $U = 23.00, z = -3.09, p = .002, r = -0.62$; however, the scores of the 14 children whose responses were consistent with the rare base rate were not different from chance.

Discussion

These results suggest that changing how 3-year-olds construe a causal mechanism from objects activating a physical machine to an agent’s preferences for the objects affected their causal inferences. Children appeared more sensitive to base rate information when the causal relation was presented as an agent’s desires. The majority of children who used the base rate information were also more likely to recognize that there was a relation between an agent’s desire and a nonobvious property of the object on the internal properties measure. This suggests that there is a relation between how children interpret a causal relation and how they observed the ambiguous data in this procedure. These inferences

Table 3
Distribution of Responses on the Internal Property Measures Across Experiments

Experiment	No. of causal responses				Statistics	
	0	1	2	3	<i>M</i>	<i>SD</i>
Experiment 2						
Rare condition ^a	7	8	8	2	1.20	0.96
Common condition ^b	6	3	2	4	1.26	1.28
Combined condition ^c	13	11	10	6	1.23	1.07
Experiment 3 ^a	7	4	8	6	1.52	1.16
Experiment 4a ^a	5	8	8	4	1.44	1.00
Experiment 4b ^b	2	5	4	4	1.67	1.05

Note. Means out of a possible 3.00.

^a *N* = 25. ^b *N* = 15. ^c *N* = 40.

might be explained by the Bayesian description we have suggested, integrating a more domain-general reasoning account with children's developing domain-specific knowledge about how to form a hypothesis space and represent the likelihood of data given those hypotheses.

However, there are several concerns with this conclusion. First, although 3-year-olds' responses were significantly different between the rare and common conditions, the overall percentage of these children who said that Object B was a blicket was 44% in the rare condition, significantly greater than the base rate (1/6 or ~17%), binomial test, $p < .005$. This was also not significantly different from chance (50%). A difficulty with this quantitative comparison comes from the nature of the test question that the children were asked. Unlike adults, who can be asked about the exact probability that Mr. Blicket likes Object B (and whose responses are quite consistent with the account on similar procedures, see Griffiths, Sobel, Tenenbaum, & Gopnik, 2009), children were asked just a yes/no question. The base rate might not exactly correspond to the percentage of children who respond in a particular manner. The more critical comparison is the difference between categorization of Object B and the two baseline objects, which have the same association with Mr. Blicket's activation. Three-year-olds treated these objects differently in the rare condition, consistent with the Bayesian description. Similarly, the fact that children treated these blocks differently across the rare and common conditions suggests some attention to the base rate.

Another concern is that the overall scores on the internal property measure were lower than those in Experiment 1. A likely explanation for this difference is experimental fatigue: Given the length of the experiment and the fact that children participated in an unrelated experiment between the two measures, they might have lost interest in the procedure and relied more on the object's perceptual similarity. To examine this possibility, in subsequent experiments, we gave children only a short break between the two measures.

Finally, the most notable concern was that the agency procedure might simply have been more interesting than the analogous blicket machine procedure. As a result, children might have brought more cognitive resources to the task (see e.g., Renninger & Wozniak, 1985). We investigated this possibility in the subsequent experiments.

Experiment 3

In Experiment 3, we considered whether the agency condition was more interesting simply because children observed the box activate spontaneously (an unfamiliar event). We did this by replicating the procedure, showing children a spontaneously activating blicket machine. If children were only more interested in the blicket machine because of its unfamiliarity, their overall level of response should be similar to that in Experiment 2. If the domain difference allowed children to understand the mechanism of the causal relation they observed, then we would expect 3-year-olds to be less sensitive to the base rate information in this procedure.

However, a critical prediction of the Bayesian account is that the children who did understand this relation would respond more consistently with the base rate data. Experiment 3 tested this by also replicating the internal properties measure. We expected the minority of children who responded given the base rate informa-

tion to score higher on this measure than the majority of children who did not appreciate the base rate information.

We did make one minor departure from Experiment 1: Here we explicitly said that the objects were causing the machine to activate and that such activation indicated that the object was a "blicket." We did this because in Experiment 2 telling the child that Mr. Blicket liked an object explicitly suggested that his desire was a function of the object, and we wanted to parallel this aspect of the procedure. At issue was whether the children recognized that this connection was due to a nonobvious property, which should have informed their inferences. Finally, only the rare condition was considered here (and in the subsequent experiments) because in the common condition, responses to Object B consistent with the base rate were identical to responses consistent with the associative information inherent in the demonstration.

Method

Participants

Twenty-five 3-year-olds (12 girls, 13 boys; age $M = 43.20$ months, $SD = 2.93$) were recruited in the same manner as the previous experiments. Four additional children were tested but not included in the analysis: 3 children failed the control trial, and 1 child's data was unusable because of experimental error. The ethnic breakdown of the sample was as follows: 20 children were White, 1 was Hispanic, and 4 were of mixed descent. No information about parental education or SES was collected.

Materials and Procedure

The same materials as in Experiment 2 were used here, except that the box was presented as a blicket machine (without the cardboard eyes). Children were introduced to the box in the same manner as in the machine condition in Experiment 1, in which it spontaneously activated while the experimenter talked to the child prior to the task. In the base rate procedure, children were told that objects that activated the machine were blickets. They were asked to categorize objects by placing them into either a blicket or a not-blicket box. In the internal properties procedure, when an object activated the machine, children were told that the object made the machine go and that it was a blicket (and the reverse when it did not have causal efficacy). Otherwise, the procedure was identical to the rare condition of Experiment 2. Children first observed that 2 of 12 objects activated the machine during training. They then received the test trial, in which Blocks A and B together made the machine go, and then Block A by itself made the machine go. They were asked to categorize whether A and B were blickets. They also received an analogous baseline trial (in which X and Y activated the machine together) and a control trial. If children failed to categorize objects correctly on the control trial, they were not included in the analysis. Children then received an internal properties measure after only a short break to avoid experimental fatigue.

Results

Results of the base rate procedure are shown in Table 2 and the internal properties procedure in Table 3. On average, children required corrective feedback on 0.56 out of the 10 training trials of

the base rate procedure. On the test trial, all children responded that Block A was a blicket. Eighteen out of the 25 children (72%) claimed Block B was as well. This was significantly greater than the 44% of children in Experiment 2, $\chi^2(1, N = 50) = 4.02, p < .05, \psi = .28$. On the baseline trial, children categorized the two objects no differently from each other and as blickets 88% of the time, not significantly different from their treatment of Object B, Wilcoxon signed rank test, $z = -1.58, ns$.

The internal properties measure was scored in the same manner as in Experiment 2. Children's performance on the three trials did not differ, Cochran's $Q(2, N = 25) = 1.17, ns$, so these data were summed together. Overall, children made causal responses in 51% of the trials ($M = 1.52, SD = 1.16$). This level of response was no different from chance (50%), $t(24) = 0.09, ns$. But again performance on the internal property measure was related to performance on the base rate measure. The 18 children who said that Block B was a blicket in the test trial of the base rate measure made causal responses 33% of the time in the internal property measure ($M = 1.00$ out of 3, $SD = 0.91$). The 7 children who said that Block B was not a blicket made a causal response on 95% of the trials ($M = 2.86$ out of 3, $SD = 0.38$), a significant difference, Mann-Whitney $U = 3.50, z = -3.74, p < .001, r = .75$. Thus, children who responded consistent with the base rate information were more likely to relate an object's causal efficacy with its insides.

Discussion

Three-year-olds' inferences about objects activating a physical machine appeared different from inferences from analogous data about an agent's preferences. Children were more likely to respond on the basis of associative information in Experiment 3 and more likely to use the base rate information in Experiment 2. But the children who did respond consistent with the base rate information seemed to recognize that there was a stable element of the objects—their insides—that could be responsible for its causal efficacy.

These data also suggest that an unfamiliar event (spontaneous activation) was insufficient to promote greater interest in the procedure, making this an unlikely explanation of the results of Experiment 2. However, these data do not rule out the possibility that there was something about the agency procedure in general that benefited 3-year-olds' inferences, not their specific understanding of desire.

Although 3-year-olds might have understood the relation between an agent's preference and an object's nonobvious properties, research in theory of mind has suggested that 3-year-olds do not understand other mental states such as thinking or knowledge (see e.g., Flavell, Flavell, & Green, 1995; Wellman, Cross, & Watson, 2001). In Experiments 4a and 4b, we considered how 3-year-olds reason about the same agency procedure used in Experiment 2, changing what mental state was represented by Mr. Blicket's activation. If children were just interested in the agency they observed, then they would be expected to reason in the same manner as in Experiment 2. However, if 3-year-olds specifically recognized the link between agent's preference and objects' non-obvious properties, then their inferences should be affected by changing that construal to a mental state that they are less likely to reason about accurately such as thought (Experiment 4a) or knowl-

edge (Experiment 4b). As in Experiment 3, we expected most children to not use the base rate information and rather to respond on an associative basis. However, as in Experiment 3, the minority of children who responded more consistently with base rates should have scored higher on the internal properties measures, suggesting they understood the relation between the particular mental state and objects' nonobvious properties.

Experiment 4a

Method

Participants

Twenty-five 3-year-olds (7 girls, 18 boys; age $M = 42.84$ months, $SD = 3.57$) were recruited in the same manner as the previous experiments. Three other children were tested but not included in the final sample because they failed the control trial. The ethnic breakdown of the sample was as follows: 24 children were White, and 1 child was Hispanic. No information about parental education or SES was collected.

Materials and Procedure

The materials were the same as in Experiment 2, except that the test boxes were labeled "Mr. Blicket is thinking about it" and "Mr. Blicket is not thinking about it." The procedure was identical to that of the rare condition of Experiment 2, except that in both the base rate and internal properties procedures, children were told that Mr. Blicket's activation indicated that he was "thinking about" the object. This change of wording was the only difference from that procedure. As in that experiment, if children failed to categorize objects appropriate on the control trial, they were not included in the final analysis. To avoid experimental fatigue, we followed the internal properties measure almost immediately after the base rate measure.

Results and Discussion

Tables 2 and 3 show the results of the base rate and internal properties measures. In the base rate measure, children required corrective feedback on an average of 0.52 out of 10 training trials. On the test trial, all children categorized Block A as something Mr. Blicket was thinking about. Eighteen out of the 25 children (72%) categorized Object B as something that Mr. Blicket was thinking about as well. This was significantly greater than the 44% of children who did so in the rare condition of Experiment 2, $\chi^2(1, N = 50) = 4.02, p < .05, \psi = .28$. Children categorized the two objects on the baseline trial as things Mr. Blicket was thinking about 100% of the time, which was significantly greater than their treatment of Object B, Wilcoxon signed rank test, $z = -2.65, p < .01, r = -0.37$. This is potentially a ceiling effect.

The internal properties measure was scored in the same manner as the previous experiments—we considered how frequently children made causal as opposed to perceptual responses. Performance among the three trials did not differ, Cochran's $Q(2, N = 25) = 0.00, ns$. Overall, children made causal responses on 1.44 (out of a possible 3, $SD = 1.00$) of the trials (48%), which was no different from chance (50%), $t(24)$ values = $-0.30, ns$. But as in Experiments 2 and 3, a different pattern emerged when perfor-

mance on the base rate measure was considered. The 7 children who categorized Object B as something Mr. Blicket was not thinking about (consistent with their use of the base rate information) made causal responses on 71% of the internal property trials ($M = 2.14$ out of 3, $SD = 1.07$). The 18 children who claimed that Mr. Blicket was thinking about Object B (inconsistent with base rate information) only made causal responses on 39% of the trials ($M = 1.17$ out of 3, $SD = 0.86$). These frequencies were significantly different from each other, Mann–Whitney $U = 28.00$, $z = -2.20$, $p < .05$, $r = -0.44$.

These data suggest that the agency manipulation alone did not affect children's inferences. Rather, 3-year-olds were more likely to appreciate mechanism information inherent to understanding another's preferences as opposed to another's thoughts. But one aspect of this procedure might have confused 3-year-olds. Unlike desire, in which the block is either liked or not liked, thinking is ubiquitous. It is implausible that Mr. Blicket could say he was not thinking about an object without actually thinking about it. Experiment 4b used a different procedure in which the agent acted like a blicket machine. We established that the agent knew about blickets and that he would say which objects had that label (indicated by his activating). As with the case of thinking, 3-year-olds have little understanding of another's knowledge (e.g., Perner, 1991; Wellman et al., 2001) and should be unlikely to recognize that the agent's knowledge about blickets was based on a nonobvious property. But the minority who respond consistently with the base rate might be more likely to do so.

Finally, a concern with the last three experiments was that preliterate children were asked to categorize objects into boxes labeled (usually) with full sentences. They might have forgotten the meanings of the boxes at some point during the procedure. Although performance during the training session and control trials suggested this was not an issue, to ensure this was not the case, we named the agent "Mr. B" and used only the letter *B* to indicate *blicket* and a blank box to indicate *not blicket*.

Experiment 4b

Method

Participants

Sixteen 3-year-olds (11 girls, 5 boys; $M = 43.25$ months, $SD = 3.00$) were recruited in the same manner as the previous experiments. Three other children were tested but not included in the final analysis because they failed the control trial. The ethnic breakdown of the sample was as follows: 13 children were White, 2 were Hispanic, and 1 was of mixed descent. No information about parental education or SES was collected.

Materials

The same materials from the previous experiments were used here, except that instead of words printed on the test boxes, one box had a *B* on it while the other was blank.

Procedure

The base rate procedure was similar to that in Experiment 2 with minor modifications. The animate detector was introduced as "Mr.

B" instead of as Mr. Blicket. After the experimenter and child "talked" to Mr. B, the block set was introduced as in the previous experiments. The experimenter then said that some of these blocks were blickets, and he asked Mr. B to show him which blocks in the set had that label. When the boxes used to categorize the objects were brought out, the box for blickets had just a *B* on it. The experimenter said, "Look, I have a box, and it has a *B* on it—*B* for blicket—let's put all the blickets in here." The other box—for not blickets—had no writing on it. The experimenter described it by saying, "Look, I have another box, and it doesn't have anything on it. Let's put all the ones that are not blickets in here." Otherwise, the base rate measure paralleled that of the rare condition in Experiment 2 (and Experiments 3 and 4a), with the test question being about whether individual objects were blickets, instead of whether Mr. Blicket liked individual objects.

The internal properties measure was the same as in Experiment 2. When Mr. B activated, children were told that "Mr. B says this one is a blicket." When he did not activate, children were told the block was not a blicket. Only a short break was taken between the two procedures.

Results and Discussion

Base Rate Measure and Internal Property Trials

Results of the base rate measure are shown in Table 2 and the internal properties measure in Table 3. On the base rate measure, children required corrective feedback on an average of 0.63 of the 10 training trials, which suggested that they understood the nature of the task. All but 1 child categorized Object A as a blicket (94% of the time). Regarding Object B, 1 child generated an ambiguous response and, as a result, was not included in the analysis.³ Twelve children (of the remaining 15, or 80%) categorized Object B as a blicket, greater than the 44% of the children in Experiment 2, Fisher's exact test, $p < .05$, $\psi = .35$. The percentage of time children categorized Object B as a blicket was no different from the percentage of time children categorized the baseline objects as blickets (90%), Wilcoxon signed rank test, $z = -1.13$, *ns*.

Performance on the three internal property trials did not differ, Cochran's $Q(2, N = 15) = 2.89$, *ns*. Children made causal responses on 56% of the trials ($M = 1.67$ out of 3, $SD = 1.05$), not different from chance (50%), $t(15) = 0.62$, *ns*. Comparing these scores in terms of performance on the base rate measure revealed that the children who categorized Object B as not a blicket (consistent with the base rate) made causal responses 89% of the time ($M = 2.67$ out of 3, $SD = 0.58$), while the children who categorized Object B as a blicket (inconsistent with the base rate) did so 47% of the time ($M = 1.42$ out of 3, $SD = 1.00$). This difference was marginally significant, Mann–Whitney $U = 5.50$, $z = -1.87$, $p = .07$, $r = -.47$.

³ In particular, this child stated that Object B was a blicket but insisted that it belonged in the non-blicket box when asked to place it one of the two boxes. This child was successful on the control trial and required no feedback on any training trial. We felt the appropriate action would be simply to exclude his data.

Comparison Among Experiments

The last four experiments examined whether children's understanding of the mechanism behind the causal relations allowed them to use base rate information to resolve ambiguous data. We do not believe that there is anything special about desire that allows children's inferences to be more consistent with the Bayesian description; only that at age 3, more children understand the mechanism behind another's desires than understand how machines work or about an agent's thinking or knowledge. It is this understanding, and not desire per se, that drives these results.

To examine this hypothesis, we combined data from children in the rare condition in Experiment 2 with data from Experiments 3, 4a, and 4b. All of these children were shown that objects with causal efficacy (desire, activating a machine, thinking, or knowledge) were rare and then were shown the same covariation data: A and B were efficacious, and then A alone was as well. These data would be ambiguous regarding the efficacy of B if children could appreciate that causal relations were deterministic. We compared the 31 children who said that Object B was not efficacious across these experiments with the 59 children who did. The former group (whose responses were consistent with the rare base rate) made causal responses on 72% of the internal property trials ($M = 2.16$ out of 3, $SD = 0.93$), while the latter group (whose responses were inconsistent with the base rate) made causal responses on only 35% of the trials ($M = 1.05$ out of 3, $SD = 0.88$). This was significantly lower, Mann-Whitney $U = 365.00$, $z = -4.84$, $p < .001$, $r = -.51$.

The data from these 90 children were also analyzed using a binary logistic regression. Whether the children categorized Object B with causal efficacy was the dependent measure. Age, experiment (i.e., how the causal relation was construed), and score on the internal properties measure were the independent variables. We first considered a model with the child's age, the experiment the child participated in, and the interaction between these variables. This model was not significant, $\chi^2(7, N = 90) = 10.06$, *ns*, and no individual factor was a significant predictor of children's treatment of Object B. We then entered the internal property score as well as all two- and three-way interactions in one block to create a full saturated model. Variables from this set were removed one at a time if they did not contribute significantly to the model (measured by the change in the $-2 \log$ likelihood of the model). Only one variable remained after these steps, that of internal property score, $\beta = -2.15$, Wald test = 20.47, $p < .001$, and the resulting overall model was significant, $\chi^2(8, N = 90) = 49.93$, $p < .001$. Across these four experiments, children's score on the internal properties measure predicted their likelihood of treating the data they observed as ambiguous and using the base rate information to resolve that ambiguity. What this suggests is that children develop mechanism knowledge across domains at different times, but when they do so, they can apply that knowledge to a more domain-general mechanism for causal inference.

General Discussion

These experiments examined whether changing the way children construed causal relations affected 3-year-olds' causal inferences. Experiment 1 found that 3-year-olds' inferences about causal relations in the physical and psychological domains dif-

ferred. Children treated an agent's desires as indicating a nonobvious property of objects but did not make the same inference when the causal relations involved objects activating a machine. Three-year-olds appeared to recognize that an agent's desires were caused by the objects, and we posited that this could be akin to what Tenenbaum and Griffiths' (2003) activation law—prior knowledge that allows children to recognize that an individual object's causal efficacy is stable.

Experiment 2 tested this hypothesis by presenting 3-year-olds with a procedure designed to examine the Bayesian description more explicitly. Three-year-olds were shown that an agent either liked many or few objects, and they then were shown covariation data that would be ambiguous if children recognized this activation law. Most 3-year-olds responded on the basis of the base rate information, particularly important in the rare condition, in which the associative information in these data suggested a different pattern of performance than recognizing the base rate. Children who reasoned according to the base rate, particularly in the rare condition, were more likely to recognize the relation between objects' causal efficacy and another nonobvious property.

In Experiments 3, 4a, and 4b, we presented children with the same covariation data as in Experiment 2 but changed the way in which the causal relations were construed. In these experiments, the majority of 3-year-olds did not appear to appreciate the rare base rate. But as in Experiment 2, the minority who did appreciate the rate also recognized that objects that shared that particular causal relation shared another nonobvious property, which could be construed as the mechanism for the causal relation. More generally, these data suggest that children integrate the covariation data they observe with their mechanism knowledge—if that knowledge suggests a deterministic relation between cause and effect, they are able to resolve ambiguous data by the base rate instead of another means, like the associative information inherent in the data. In general, these experiments supported a model of children's causal learning that integrates their developing domain-specific knowledge with a more domain-general mechanism best described by Bayesian inference.

It is important to note that these experiments do not tap every aspect of physical or psychological knowledge. There is clearly more to the physical domain than understanding whether objects activate a machine and more to the psychological domain than understanding the mechanisms behind another's desires, thoughts, or knowledge. An open question is whether there are other pieces of physical or psychological knowledge or knowledge in other domains that 3-year-olds (or potentially even younger children) recognize as having stable mechanisms. If so, we might expect to see similar performance on our base rate measure. The present experiments suggest that as children begin to understand specific aspects of causal knowledge in a domain, their understanding is coherent and can be used to interpret data they observe. The Bayesian approach makes a case for how this is done for the particular inference we considered.

Therefore, we wish to conclude that the present data are not simply due to a general facilitative effect of introducing domain differences. This conclusion is supported by the logistic regression described in the last experiment. The experiment that the children were in (and hence what construal of the causal relation they were shown) did not significantly predict their use of the base rate information. What did predict their inferences was whether they

recognized that the causal relation with which they were presented (regardless of construal) indicated a potential mechanism. A strength of the Bayesian description is that it predicts why this facilitation should take place: if children understand that causes relate to their effects deterministically, then the observed data are ambiguous and the base rate is relevant.

Two concerns, however, do remain. First, how do children make causal inferences when they do not appreciate the mechanism behind the data they observe? For example, it is unclear whether 3-year-olds who do not use the base rate information believe that objects do not always activate the machine (which would result in stating that Object B has efficacy) or whether they have no understanding and thus rely on the associative information present (which would result in the same response). We are agnostic on this issue, but the latter does seem a simpler explanation.

Second, we have only considered a particular causal inference about ambiguity across the last four experiments. Our goal was to equate the data children observed and vary only the way in which causes related to effects was construed. As such, we kept the covariation data that children were shown constant and consistent with a previous study (Sobel et al., 2004). The Bayesian description presented here has been extended to other causal reasoning problems and patterns of covariation data, which preschoolers appear to solve consistent with the model's predictions (see e.g., Griffiths et al., 2008; Schulz, Bonawitz, & Griffiths, 2007; Schulz & Gopnik, 2004).

To conclude, these experiments were consistent with a computational-level description of children's causal reasoning that is best described by Bayesian inference. This leaves open the question as to the process (or algorithmic-level description) by which children make these inferences. It seems unlikely to us that children form specific hypotheses, particularly given data that suggest preschoolers have poor scientific reasoning abilities (e.g., Kuhn, 1989). It is also unlikely that young children can explicitly compute Bayesian statistics or even understand these probabilistic concepts (much research suggests the opposite, e.g., Piaget & Inhelder, 1975). But children *are* responding consistently with their making complex causal inference. Some have explained this by appealing to other computational accounts, like parallel distributed processing (PDP) models, to describe this algorithmic level (e.g., McClelland & Thompson, 2007; Rogers & McClelland, 2004). Others have suggested that a process account of human inference is based in logical reasoning (e.g., Lovibond et al., 2003). One could also imagine children's developing executive function as important to the means by which children engage in causal inference (e.g., Frye, Zelazo, Brooks, & Samuels, 1996). Regardless, whatever the process by which children engage in causal inferences, it must be influenced by their domain-specific knowledge.

References

- Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *54*, 299–352.
- 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 Hirschfeld, Lawrence A. & Gelman, S. A. (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169–200). New York: Cambridge University Press.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405.
- Fawcett, C., & Markson, L. (2005, August). *Developing ideas about other's preferences*. Poster session presented at the biennial meeting of the Cognitive Development Society, San Diego, CA.
- Flavell, J. H., Green, F. L., & Flavell, E. R. (1995). Young children's knowledge about thinking. *Monographs of the Society for Research in Child Development*, *60*(1), Series No. 243.
- Frye, D., Zelazo, P. D., Brooks, P. J., & Samuels, M. C. (1996). Inference and action in early causal reasoning. *Developmental Psychology*, *32*, 120–131.
- Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3–32.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
- Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development*, *62*, 98–110.
- 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 co-variation. *Developmental Psychology*, *37*, 620–629.
- Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2009). *Bayesian reasoning in adults' and children's causal inferences*. Manuscript submitted for publication.
- Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*, 285–386.
- Inagaki, K., & Hatano, G. (1993). Young children's understanding of the mind–body distinction. *Child Development*, *64*, 1534–1549.
- Johnson, S. C., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science*, *1*, 233–238.
- Kalish, C. W. (1996). Preschoolers' understanding of germs as invisible mechanisms. *Cognitive Development*, *11*, 83–106.
- Kuhlmeier, V. A., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*, 402–408.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*, 674–689.
- Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*, 265–288.
- Lovibond, P. F., Been, S., Mitchell, C. J., Bouton, M. E., & Frohardt, R. (2003). Forward and backward blocking of causal judgment is enhanced by additivity of effect magnitude. *Memory & Cognition*, *31*, 133–142.
- Marr, D. (1982). *Vision*. New York: Henry Holt.
- McClelland, J. L., & Thompson, R. M. (2007). Using domain-general principles to explain children's causal reasoning abilities. *Developmental Science*, *10*, 333–356.
- Meltzoff, A. N. (1995). Understanding the intentions of others: Reenactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838–850.
- Michotte, A. (1962). *Causalite, permanence, et realite phenomenales* [Causality, permanence, and reality]. Oxford, England: Publications Universitaires de Louvain.
- Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. *Cognitive Development*, *5*, 193–207.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York: W. W. Norton.
- Renninger, K. A., & Wozniak, R. H. (1985). Effect of interest on atten-

- tional shift recognition and recall in young children. *Developmental Psychology*, 21, 624–632.
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, 33, 12–21.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Rosenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26, 521–562.
- Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, 16, 995–1001.
- Schulz, L., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared give you a tummy-ache? Naïve theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology*, 43, 1124–1139.
- Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology*, 40, 162–176.
- Shanks, D. R. (1995). Is human learning rational? *Quarterly Journal of Experimental Psychology A*, 48, 257–279.
- Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, 47(1), Serial No. 194.
- Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: The development of causal reasoning in toddlers and infants. *Developmental Psychology*, 42, 1103–1115.
- Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, 28, 303–333.
- Sobel, D. M., Yoachim, C. M., Gopnik, A., Meltzoff, A. N., & Blumenthal, E. J. (2007). The blicket within: Preschoolers' inferences about insides and causes. *Journal of Cognition and Development*, 8, 159–182.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99, 605–632.
- Tenenbaum, J. B., & Griffiths, T. L. (2003). Theory-based causal inference. In S. Thrun, L. K. Saul, & B. Schölkopf, *Advances in neural information processing systems 16: Proceedings of the 2003 conference*. Cambridge, MA: MIT Press.
- Tenenbaum, J. B., Griffiths, T. L., & 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.
- Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.
- Wellman, H. M., Cross, D., & Watson, J. K. (2001). A meta-analysis of theory of mind: The truth about false belief. *Child Development*, 72, 655–684.
- Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, 35, 245–275.

Received June 25, 2007

Revision received November 13, 2008

Accepted November 24, 2008 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.