

Teaching three-and-a-half-year-olds to revise their beliefs given ambiguous evidence

Elizabeth B. Bonawitz¹, Adina Fischer², & Laura Schulz³

¹ Correspondence to: 5415 Tolman Hall, Department of Psychology, University of California, Berkeley, Berkeley CA, {liz_b@berkeley.edu}

² Dartmouth College, Hanover, NH 03755

³ Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract

Previous research suggests that three-year-olds fail to learn from statistical data when their prior beliefs conflict with evidence. Are children's beliefs entrenched in their folk theories, or can preschoolers rationally update their beliefs? Motivated by a Bayesian account, we conducted a training study to investigate this question. Children (45 months) who failed to endorse a statistically more probable (but a priori unlikely) cause following ambiguous evidence were assigned to a *Statistical Reasoning* training, one of two *Prior Belief* trainings (*Base Rates*, *Mechanisms*), or a *Control* condition. Relative to the *Control*, children in the trainings were more likely to endorse the a priori unlikely variable on a free-explanation task. Critically, children in the *Statistical Reasoning* condition passed this task, even though their only information about the belief-violating variable came from ambiguous evidence. This suggests that statistical reasoning training improves preschoolers' ability to learn even from data inconsistent with their prior beliefs.

Keywords: Causal learning; Ambiguous evidence; Training study, Psychosomatic illness.

Teaching three-and-a-half-year-olds to revise their beliefs given ambiguous evidence

The view that children's causal commitments take the form of naïve theories has been influential in developmental psychology for several decades (Carey, 1985; Harris, German, & Mills, 1996; Hickling & Wellman, 2001; Gopnik, 1988; Gopnik & Meltzoff, 1997; Keil, 1989; Perner, 1991; Sobel, 2004; Wellman, 1990; Wellman, Hickling, & Schult, 1997). This view, the *theory theory*, maintains that there is a dynamic relationship between children's naïve theories and evidence such that children's causal beliefs affect their interpretation of evidence, and evidence enables children to revise their beliefs (e.g., Gopnik & Meltzoff, 1997). However, previous studies looking at the relationship between patterns of evidence and children's folk theories have generated contradictory results, with some studies suggesting that children privilege plausible causal mechanisms over statistical evidence (e.g., Shultz, 1982) and others suggesting that children learn theory-violating causal relations as readily as theory-consistent ones (Kushnir & Gopnik, 2007; Schulz & Gopnik, 2004). To the extent that children's causal judgments reflect an interaction between their naïve theories and patterns of evidence, such conflicting findings are perhaps not surprising: on any given task, children's inferences could depend on the strength of their initial beliefs, the strength of the data, and their ability to integrate the two.

Recently, Bayesian analyses have offered a formal account of how learners might integrate statistical evidence with constraints from naïve theories (Kemp, Perfors, & Tenenbaum, 2007; Tenenbaum, Griffiths, & Kemp, 2006). In Bayesian inference, a learner observes some data, D , and seeks to evaluate a hypothesis, h , about the process that produced the data. The learner's evaluation of the hypothesis given the data, $P(h|D)$,

depends on the learner's prior belief in the probability of the hypothesis, $P(h)$, and the likelihood of the hypothesis, $P(D|h)$, the probability that the data would have been observed if the hypothesis were true. Bayes' law specifies that $P(h|D) \propto P(h)P(D|h)$. Bayesian inference thus formalizes the claim that evidence and prior beliefs interact to affect children's causal judgments.

To our knowledge, however, only one previous study has directly tested whether, consistent with the predictions of a Bayesian inference model, children can use probabilistic evidence to revise their domain-specific beliefs (Schulz, Bonawitz, & Griffiths, 2007). In that study, children were provided with ambiguous data both in contexts where they had strong prior beliefs and ones where they did not. In considering a context in which children had strong prior beliefs, the researchers focused on psychosomatic causality because studies suggest that children initially draw a sharp distinction between mental phenomena and bodily/physical phenomena (Bloom, 2004; Carey, 1985; Estes, Wellman, & Woolley, 1989; Hatano & Inagaki, 1994; Notaro, Gelman, & Zimmerman, 2001; Wellman & Estes, 1986). Thus, although older children accept psychogenic events, preschoolers discount the plausibility of psychosomatic phenomena (e.g., they deny that feeling embarrassed can make one blush or that feeling frustrated can cause a headache; Notaro et al., 2001). The fact that children change their minds about psychosomatic causality suggests that this may be a viable context in which to investigate children's ability to revise their causal beliefs through evidence.

In the Schulz et al. (2007) study, preschoolers were read two books in which two candidate causes co-occurred with an effect. Evidence was presented in the form $AB \rightarrow E$; $CA \rightarrow E$; $AD \rightarrow E$, etc. After receiving this evidence, children were asked to identify the cause of the effect on a new trial. Although it was possible that B, C, D, etc. were each

independent causes of the effect, it was more probable that the recurring cause, A, was the actual cause. In one book (the Within Domain book), all the causes and the effect were from a single domain (physical/bodily events); in the other book (the Cross Domains book), the recurring cause, A, came from a different domain (A was a psychological cause of a bodily effect). Thus, in the Cross Domains book, there was a conflict between the statistical evidence and children's prior beliefs.

Consistent with the prescriptions of Bayesian inference, kindergartners and older preschoolers (50-70 months; mean: 60 months) correctly inferred that A was the cause in both cases but were more likely to identify A as the cause in the Within Domain book than the Cross Domains book. Critically, younger preschoolers (42-48 months; mean: 45 months) did not. Although the three-and-a-half-year-olds readily identified cause A as the target cause in the Within Domain book (indeed, they were indistinguishable from the older children), they failed to learn at all in the Cross Domains book: they consistently chose the within-domain cause.

The dramatic contrast between the three-and-a-half-year-olds' impressive reasoning about statistical evidence in neutral contexts (near ceiling), and their poor reasoning (near floor) in belief-violating contexts, leaves open the possibility that very young children might be able to use the relative probabilities of events to distinguish candidate causes, but that children's prior beliefs about plausible causal relations restrict the kinds of evidence they are willing to consider in the first place.

Note that with the exception of the study cited above (see also Kushnir & Gopnik, 2007), virtually all of the previous work showing that children can use statistical data to revise their prior beliefs has used unambiguous deterministic evidence

and focused on older children (four to six-year-olds; e.g., Schulz & Gopnik, 2004; Sobel, 2006). Other work shows that children as young as three can learn from probabilistic and ambiguous evidence and both explore and seek explanations for such evidence (Bonawitz, Lim, & Schulz, 2007; Legare, Gelman, & Wellman, 2010; Schulz, Goodman, Tenenbaum, & Jenkins, 2008; Schulz, Hooppell, & Jenkins, 2008; Schulz & Sommerville, 2006; Sobel & Munro, 2009). However, if anything, these studies show that children try to explain away anomalous evidence (Legare et al., 2010; Schulz & Sommerville, 2006); no previous work has suggested that young children can use ambiguous evidence to revise their beliefs.

There are at least two reasons why it is important to look at whether very young children can revise their beliefs given ambiguous evidence. The first is that, in the real world, children are much more likely to have access to ambiguous than unambiguous evidence for causal relationships (i.e., because children may fail to observe or encode relevant evidence, because causal variables are often confounded, etc.). The theory theory proposes that evidence-based learning occurs from the earliest stages of development; if the available evidence is often ambiguous, then investigating young children's abilities to reason from ambiguous evidence is a critical test of this account. The second reason to look at children's ability to revise their beliefs from ambiguous data is that Bayesian inference has been proposed as a formal model underlying the theory theory account (see e.g., Gopnik & Schulz, 2007). This computational account presumes that all rational learners can update their beliefs from probabilistic data. However, to date, studies have shown that although four and five-year-olds can revise their beliefs from ambiguous data, three-year-olds do not (Kushnir & Gopnik, 2007;

Schulz et al., 2007). Critically, the children in these studies have been exposed to very small amounts of evidence (e.g., a few trials of data at a single session). A training study, giving children more extensive exposure to evidence, provides a more robust test of children's inferential abilities. Here we look at whether evidence-based teaching is effective in helping young preschoolers revise their causal beliefs given ambiguous data. Inspired by rational Bayesian inference accounts, we introduce three training conditions intended to independently affect children's statistical reasoning and their prior beliefs.

In one training condition (the *Statistical Reasoning* condition), we teach children to draw accurate inferences when multiple causal relations are possible, but one is more probable than the others (i.e., we read children stories in which evidence is presented in the form $AB \rightarrow E$; $CA \rightarrow E$; $AD \rightarrow E$, etc. and teach children that A is the most likely cause of the effect, E). This condition was motivated in part by the finding that young three-year-olds (3;0 - 3;5) fail to learn from such evidence even in a theory-neutral (Within Domain) condition (Schulz et al., 2007). Given that even the neutral task was challenging for three-year-olds, three-and-a-half-year-olds might not have been able to handle the increased difficulty posed by a conflict with prior beliefs. If giving three-and-a-half-year-olds practice reasoning from evidence in *theory-neutral* contexts improves their ability to draw accurate inferences from *theory-violating* evidence, this would suggest that even young preschoolers can use formal properties of evidence, not merely to distinguish plausible causes, but also to draw inferences at a more abstract level: inferring the causal status of previously unrecognized causal variables.

As noted, however, the Bayesian analysis suggests that children's causal judgments depend not only on the strength of the evidence but also on their prior

beliefs. Thus in the other two training conditions, we intervene on children's causal beliefs. How might evidence-based teaching help children revise their prior beliefs about psychosomatic causality? Arguably, young children have relatively limited exposure to psychosomatic events. If children believe the base rate of psychosomatic causality is low, then they might (rationally) resist accepting a psychological cause as the most probable explanation of a bodily effect. Thus, one approach to intervening on children's prior beliefs is to manipulate their perception of the frequency of psychosomatic events. We will call this the *Prior Beliefs Base Rate* condition.

Another reason children might resist psychosomatic causality is because they do not understand *how* psychological states affect bodily states. Research suggests that both adults and children are more willing to accept causal relations for which they can imagine plausible mechanisms (e.g. Ahn, Kalish, Medin, & Gelman, 1995; Shultz, 1982). Thus, we might increase children's acceptance of psychosomatic causality by offering an explanation of how emotional states might cause bodily outcomes. We will call this the *Prior Beliefs Mechanism* condition.

If children who initially fail to endorse psychosomatic causality succeed after any or all of the three training conditions, this would suggest that well before children begin formal education, rational inductive inference mechanisms support children's ability, not only to reason about theory-neutral causal relations, but also their ability to reason about a priori unlikely events. Note that children might succeed in some of the training conditions and not others (e.g., intervening on children's prior beliefs might be helpful, but intervening on their statistical reasoning ability might not). However, given that the Bayesian analysis suggests that both factors play a role, we predict that all three

manipulations should independently improve children's performance. By contrast, if three-year-olds' causal beliefs are relatively entrenched (e.g., because they are constrained by core knowledge in domains like naïve physics and naïve psychology), then we would expect children to be resistant to belief revision and perform no better than children in a control condition.

Training Study

We designed a two-week training study to investigate these accounts. Because we were interested in children's ability to reason about theory-violating evidence, children were included in the study only if they initially endorsed the plausible within-domain cause rather than the statistically likely, but theory-violating, cause in a pretest book (identical to the Cross Domains book used in Schulz et al., 2007). Eighty children were assigned to one of four conditions: a *Statistical Reasoning* training, a *Prior Belief Base Rates* training, a *Prior Belief Mechanisms* training, and a *Control* condition. At the final session, children were tested on two dependent measures. First, they were again read a Cross Domains storybook (formally identical to the initial book but with different specific stimuli). However, given that previous research suggests that preschoolers tend to vary their responses when asked versions of the same question twice (e.g. Memon, Cronin, Eaves, & Bull, 1993; Poole & White, 1991), we believed that this measure might not be the most sensitive index of children's learning. Thus, our second dependent measure of interest was a free explanation task, adapted from Schulz et al. (2007), originally given to four and five-year-olds. The free explanation measure provides a strong test of children's learning in that children have to transfer their knowledge to a novel task.

In the free explanation task, children were told about a puppy dog who was scared about the first day of school and had a tummy ache; children were asked to explain why the puppy had a tummy ache. Researchers in the earlier study found that at baseline four and five-year-olds ignored the only variable mentioned in the story (being scared), and strikingly, they invented their own domain-appropriate explanations instead (e.g., “because he fell on his stomach”; “because he ate too much food”); children who had first been exposed to the evidence in the Cross Domains book adopted the psychosomatic explanation. If the training conditions support three-and-a-half-year-olds’ ability to learn the target causal relation, they should be able to transfer their learning and explain the bodily event using the psychological explanation.

Methods and Design

Participants

Eighty children (mean age: 45 months; range: 39-48 months; 54% girls) were recruited from preschools in a metropolitan area. An experimenter met individually with each child for four 20-minute sessions over a period of two weeks. No two sessions were on consecutive days. Most of the children were white and middle class but a range of ethnicities resembling the diversity of the population was represented.

To ensure that participants entering the training did not already endorse psychosomatic events, children were given an initial Cross Domains test book; those who passed were dropped from the study and replaced. Replicating Schulz et al. (2007), 82% of the three-and-a-half-year-olds tested on the initial Cross Domains storybook failed the task (i.e., chose the theory-consistent rather than the statistically probable cause), and thus met the inclusion criteria for the study. Children were randomly assigned to a

Statistical Reasoning Training condition, *Prior Belief Base Rates Training* condition, a *Prior Belief Mechanisms Training* condition, or a *Control* condition (20 children per condition). There were no age differences among the four conditions ($F(3, 76) = 1.48, p = ns$).

Materials

Two *Cross Domains* books and a *Free Explanation* test book were used. Additionally, five different training books were used in each of the four conditions (*Statistical Reasoning*, *Prior Belief Base Rate*, *Prior Belief Mechanisms*, and *Control*), for a total of 20 training books. The training books were each approximately 20 pages long and had approximately 9 words per page. (See Figure 1.)

Cross Domains books. One book was used to see if children met the initial inclusion criteria, and one was used as the first dependent measure on the final day, preceding the *Free Explanation* Test book. The books were identical except for details of the stimuli. In each book, a character (Bunny or Beaver) ate a different food, experienced a recurring psychological cause (feeling worried; feeling scared), and a recurring biological effect (belly ache; tummy hurting) each morning of a seven-day week. Each afternoon, the character ate two different foods and felt fine. At the end of the story, children were asked a forced choice question about the events of that morning: “Why does (Bunny’s, Beaver’s) (belly ache? tummy hurt)? Is it because of (feeling worried, feeling scared) or because of eating (the cornbread, the sandwich)?” The order of events (psychological or food) was counterbalanced throughout.

Training books. Five books, each involving unique characters and candidate relations, were used in each training condition.

Statistical Reasoning Training. In each book, a character experienced a pair of candidate causes (one recurring and one varying each day) and a consistent effect in a format identical to the Cross Domains books ($AB \rightarrow E$; $CA \rightarrow E$; $AD \rightarrow E$, etc.). In each book all the variables were drawn from a single domain; no domains were psychological. At the end of each story, children were given a forced choice between causal variables (e.g.: “Why does Bambi have itchy spots? Is it because of running in the cattails or running in the garden?”)

Prior Belief Base Rate Training. Each book showed ten characters in a classroom. All ten characters experienced the same emotion (e.g. boredom waiting for a hamster to do a trick). Eight of the ten characters had a bodily reaction (e.g. Sue gets sleepy; Charles gets sleepy; Josh does not get sleepy). At the end of the book children were given a forced choice question asking whether the bodily reaction to the psychological emotion happened to very many or very few characters in the story (e.g., “Can you remind me: did very many students get sleepy or did very few students get sleepy?”).

Prior Belief Mechanisms Training. Each book explained that a particular psychological state could generate bodily effects and offered a brief account of how this might happen (e.g. “When Peter feels embarrassed, his brain makes different things happen to his body...his cheeks turn pink and he starts to blush. That’s because Peter’s brain changes the way energy moves through his body and can send energy to his cheeks.”) At the end of each book, children were asked to repeat the explanation for the bodily outcomes in the books (e.g., “Can you explain to me: what made Peter blush?”).

Control. The control books told a story about a character who experienced a recurring psychological state throughout the day (e.g. “Tom is excited because today is

his birthday. In the morning, Tom’s mom gives him a present. Tom is very excited to open his first present.”). To match the level of engagement in the other training conditions, children were asked memory questions at the end of each story (e.g. “Can you explain to me: why did Tom have to take off the ribbon?”).

Free Explanation test book. This book read in its entirety: “This is Puppy. Puppy is nervous because it’s his first day of school. Oh, oh! Puppy’s stomach hurts!” Children were asked: “Why does Puppy’s stomach hurt?”

Procedure

Children were tested individually in a quiet room at their daycare. Participants were first tested on one of the two Cross Domains books (particular book counterbalanced between children; the other book was then used on the final day).

Children who met the criteria for the training study were read the first book from their assigned condition (see Figure 1). The experimenter then met with the child three more times over the course of two weeks. On each of the second and third visits, children were read the two books appropriate to their training condition (Books 2 & 3 on Day 2; and Books 4 & 5 on Day 3). The experimenter gave feedback if the child answered incorrectly during the training sessions (i.e., in the *Statistical Reasoning* training, the experimenter pointed to the recurring variable and showed the child how it occurred each day along with the effect; in the *Prior Beliefs Base Rates* training, the experimenter pointed to the number of children with the bodily response and observed that it was “very many” rather than “very few”; in the *Prior Beliefs Mechanism* training, the experimenter repeated the explanation for the bodily effect; in the *Control* condition, the children were reminded of the correct information). On the final day (Day 4) the children were first

tested on the final Cross Domains storybook and then on the Free Explanation book (the order was fixed so that if children learned from the Cross Domains book, they could use the evidence and transfer their knowledge for the Free Explanation question). No feedback was given on the final day. Trainers read all (and only) the words presented on the book and pointed to the pictures on the page as they read. In the few instances that children were visibly distracted or interrupted with a comment or question, the trainer refocused the child by suggesting that they should continue to read the book together.

Preliminary analysis

Training books. Across the training period, children's performance on the training books improved. In all three training conditions, children were more likely to answer the prompts at the end of the training books correctly on the last day's training book than on the first day's (first book 45%, last book 75% in *Statistical Reasoning: McNemar* ($n = 20$), $p < .05$; first book 50%, last book 75% in *Prior Belief Base Rates*: ($n = 20$), $p = .06$); first book 35%, last book 85% in *Prior Belief Mechanism*: ($n = 20$), $p < .01$). This suggests that the training itself was implemented successfully.

Results

Cross Domains book (final day)

Responses on the final Cross Domains book were coded as appealing to the recurring psychosomatic cause or to the alternative domain-appropriate cause (i.e., the particular food). Compared to their responses on the original test book (at floor due to the inclusion criteria), children were more likely to appeal to psychosomatic causes in all conditions (*Statistical Reasoning: McNemar* ($n = 20$) = 35%, $p < .01$; *Prior Belief Base Rates*: ($n = 20$) = 45%, $p < .01$; *Prior Belief Mechanism*: ($n = 20$) = 45%, $p < .01$;

Control: ($n = 20$) = 30%, $p < .05$). There were no significant differences among conditions. As noted, however, performance on the second Cross Domains book is arguably neither the most sensitive nor the strongest measure of learning: it is relatively insensitive because preschoolers are vulnerable to varying their responses in response to repeated questioning (Memon et al., 1993; Poole & White, 1991) and relatively weak because learning from one Cross Domains book to another does not require any degree of generalization. Children's ability to transfer their learning to the open-ended explanation task is thus both a more reliable and more stringent assessment of whether the training affected children's beliefs about the plausibility of psychological variables as potential causal explanations of bodily effects.

Free Explanation book

Children's responses on the Free Explanation book were coded as appealing to the target psychological cause in the story (e.g. feeling nervous; thinking about school), to external domain-appropriate bodily causes not mentioned in the story (e.g., "eating too much food", "bumping his belly") or other. Two children (one in the *Prior Belief Base Rates* Training and one in the *Control* condition) responded, "I don't know". Otherwise, children's responses fell uniquely into the psychogenic or bodily category.

There were no significant differences among the three training conditions (χ^2 (2, $n = 60$) = .93, $p = ns$). As predicted, however, children were significantly more likely to appeal to psychological explanations across all three training conditions than the *Control* condition (χ^2 (2, $n = 80$) = 6.94, $p < .01$): 50% of the children appealed to the psychological cause in the *Statistical Reasoning* condition; 40% in the *Prior Belief Base Rate* condition, and 55% in the *Prior Belief Mechanism* training, but only 15% of

children did so in the *Control* condition (see Figure 2). Note that few children in the *Control* condition invoked the psychological variable on the free explanation task despite the fact that they read stories that repeatedly referred to a character experiencing psychological states (Tom is excited, Lisa is angry, Erin is sleepy, etc.).

Comparing individual training conditions with the *Control* condition, significantly more children appealed to psychological explanations in the *Statistical Reasoning* condition, ($\chi^2(1, n = 40) = 5.58, p < .05$) and the *Prior Belief Mechanism* condition ($\chi^2(1, n = 40) = 7.03, p < .01$) than the *Control* condition and there was a trend for children to be more likely to appeal to psychological explanations in the *Prior Belief Base Rates* condition than in the *Control* condition, ($\chi^2(1, n = 40) = 3.14, p = .08$). Interestingly, the results in the *Mechanism* condition were more robust than those in the *Base Rates* condition, consistent with the claim that children's prior beliefs may be particularly sensitive to information about plausible causal mechanisms (Ahn et al. 1995; Shultz, 1982). However, the fact that children also robustly succeeded in the *Statistical Reasoning* training suggests that, consistent with the Bayesian analysis, both prior beliefs and the strength of the evidence independently affect children's causal inferences.

Discussion

These results suggest that children as young as three-and-a-half are not unduly entrenched in their prior causal commitments; very young preschoolers rationally update their beliefs from evidence. Following a brief training, involving only five storybooks, children were able to invoke previously unrecognized causal variables as explanations for events.

What can our training study tell us about the role of statistical inference and prior beliefs in children's reasoning about theory-violating evidence? First, consider the implications of the *Statistical Reasoning* training. The children in this condition were given no more information about psychosomatic events than children in the *Control* condition -- and indeed, were given *less* exposure than children in the *Control* condition to psychological variables in general (psychological variables were never mentioned during their training). Nonetheless, children in the *Statistical Reasoning* training were more likely than children in the *Control* condition to adopt psychosomatic explanations. Thus, critically, although all children saw identical data in the two Cross Domains books, the children taught to reason about statistical evidence were better able than children in the *Control* condition to bring this evidence to bear on the explanation task.

Previous studies have shown that young children are sensitive to inconsistent and ambiguous data and can generate plausible causal explanations for anomalous evidence (e.g., Legare, et al., 2010; Schulz, Standing, & Bonawitz, 2008; Schulz et al., 2008; Schulz & Sommerville, 2006). Critically, however, these studies have largely shown that children try to dismiss ambiguous evidence rather than change their causal beliefs. When shown for instance that an object expected to activate a toy fails to do so, children claim that "the object is broken" or was "set on the wrong side" (Legare et al., 2010). By contrast, the current research shows that children as young as three-and-a-half can use ambiguous statistical data to attribute causal power to theory-violating causal variables.

The success of the *Prior Belief* conditions suggests that either increasing children's perception of the base rate of a target causal relation or increasing their understanding of the target causal mechanism may increase children's willingness to

appeal to causal relations that they previously dismissed. Although the explanations given to children in the *Mechanism* training were shallow, akin to what researchers have called “coarse sketches” (Keil, 2003; p. 371), such explanations improved children’s willingness to invoke psychological variables as causes of bodily events, whereas simple repeated exposure to psychological variables (as in the *Control* condition) did not. Consistent with other research, this suggests that even lean representations of causal mechanisms can support accurate causal judgments (see Keil, 2003; Rosenblit & Keil, 2002).

However, these results fall short of demonstrating that manipulating children’s prior beliefs improves their ability to learn from (erstwhile) implausible data. Children in the *Prior Belief* conditions might have been better able to learn from the statistical evidence in the Cross Domains books and bring this evidence to bear on the free explanation task, but it is also possible that these children did not generalize from the ambiguous evidence in the Cross Domains books; instead they may have learned about psychosomatic causation from the training books directly. That is, the exposure to the training books by itself might have made children more willing to appeal to psychosomatic causes in the free explanation task (independent of the evidence in the Cross Domains book). Further research is needed to isolate the effect of rational constraints from the effects of other task demands to see whether changing prior knowledge can change the statistical inferences of even very young children.

Finally, although this study suggests that three-and-a-half-year-olds can use rational inductive inference to revise their beliefs, we do not know at what level of abstraction such revision occurred. Previous work (Schulz et al., 2007) suggested that

four and five-year-olds changed their inferences only at a quite specific level (e.g., entertaining the possibility that worrying could cause tummy aches) rather than at the level of more abstract theories (revising their beliefs about psychogenic causality generally). In light of that, it is noteworthy that three-and-a-half-year-olds in the *Prior Beliefs* training conditions in the current study were able to generalize from other instances of psychosomatic causation to the test exemplar. This suggests that at least some abstract inferences were enabled by those training conditions. Further research might establish the degree to which different interventions transform children's ability to reason about theory-violating evidence at different levels of abstraction.

Collectively, however, these results suggest the importance of rational constraints on causal learning in early childhood. Moreover, our study suggests the malleability of some of these constraints. From even limited instruction in causal mechanisms, base rates, and statistical data, young preschoolers can learn to change their minds.

References

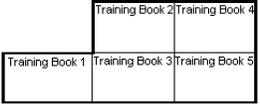
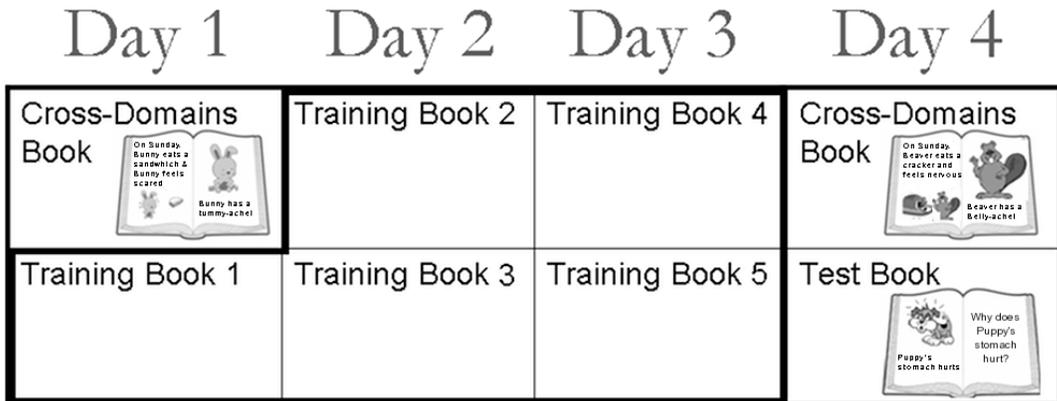
- Ahn, W., Kalish, C.W., Medin, D.L., & Gelman, S.A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *54*, 299–352.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press/Bradford Books.
- Bloom, P. (2004). *Descartes' Baby*. New York: Basic Books.
- Bonawitz, E.B., Lim, S., & Schulz, L.E. (2007). Weighing the evidence: Children's theories of balance affect play. *Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society*. Nashville, Tennessee.
- Estes, D., Wellman, H. M., & Woolley, J. (1989). Children's understanding of mental phenomena. In H. Reese (Ed.), *Advances in child development and behavior* (pp. 41-86). New York: Academic Press.
- Gopnik, A. (1988). Conceptual and semantic development as theory change. *Mind and Language*, *3*(3), 197-217.
- Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts and theories*. Cambridge, MA: MIT Press.
- Gopnik, A. & Schulz, L. E. (2007). *Causal learning: Psychology, Philosophy, and Computation*. New York: Oxford University Press.
- Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*, 233-259.
- Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition*, *50*, 171–188.

- Hickling, A. K., & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology*, 37, 668-683.
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Keil, F.C. (2003). Folk Science: Coarse interpretations of a complex reality. *Trends in Cognitive Science*. 7, 368-373.
- Kemp, C., Perfors, A. & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, 10, 307-321.
- Kushnir & Gopnik (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology*, 43, 186-196.
- Legare, C.H., Gelman, S.A., & Wellman, H.W. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development*, 81, 929-944.
- Memon, A., Cronin, O., Eaves, R., & Bull, R. (1993) The cognitive interview and child witnesses. In G.M. Stephenson & N.K. Clark (Eds), *Children, Evidence, and Procedure. Issues in Criminological and Legal Psychology*. No. 20. Leicester, UK: British Psychological Society.
- Notaro, P.C., Gelman, S., & Zimmerman, M.A. (2001) Children's understanding of psychogenic bodily reactions. *Child Development* 72, 444-459.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

- Poole, D.A. & White, L.T. (1991) Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, 27, 975-986.
- Rozenblit, L.R. and Keil, F.C. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science*, 26, 521-562.
- Schulz, L.E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology*, 40, 162-176.
- Schulz, L., Bonawitz, E.B., & Griffiths, T. (2007). Can being scared make your tummy ache? Naïve theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology*, Sep Vol 43, 1124-1139.
- Schulz, L.E., Goodman, N.D., Tenenbaum, J.B., & Jenkins, C.A. (2008). Going beyond the evidence: Abstract laws and preschoolers' responses to anomalous data. *Cognition*, 109, 211-223.
- Schulz, L.E., Hoopell, K., & Jenkins, A. (2008). Judicious Imitation: Young children imitate deterministic actions exactly, stochastic actions more variably. *Child Development*, 79, 395-410.
- Schulz, L. E. & Sommerville, J. (2006). God does not play dice: Causal determinism and children's inferences about unobserved causes. *Child Development*, 77, 427-442.
- Schulz, L.E., Standing, H., & Bonawitz, E.B. (2008). Word, thought and deed: The role of object labels in children's inductive inferences and exploratory play. *Developmental Psychology*, 44, 1266-1276.
- Shultz, T. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, 47.

- Sobel, D. M. (2004). Exploring the coherence of young children's explanatory abilities: Evidence from generating counterfactuals. *British Journal of Developmental Psychology, 22*, 37-58.
- Sobel, D. (2006). When Mr. Blicket wants it, children are Bayesian. *Proceedings of the Cognitive Science Society*. 810-816.
- Sobel, D. M., & Munro, S. A. (2009). Domain generality and specificity in children's causal inferences about ambiguous data. *Developmental Psychology, 45*, 511-524.
- Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Science, 10*, 309-318.
- Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.
- Wellman, H. M., & Estes, D. (1986). Early understanding of mental entities: a reexamination of childhood realism. *Child Development, 57*, 910-923.
- Wellman, H. M., Hickling, A. K., & Schult, C. A. (1997). Young children's psychological, physical, and biological explanations. In H. M. Wellman & K. Inagaki (Eds.), *The emergence of core domains of thought: Children's reasoning about physical, psychological, and biological phenomena. New directions for child development No. 75* (pp. 7-25). San Francisco, CA: Jossey-Bass/Pfeiffer.

Figure 1: Study design and sample pages from the training books



Statistical Reasoning Training Books

(Example Book)

"Why does Bambi have itchy spots? Is it because of running in the garden or running in the cattails?"

Prior Belief Mechanism Training Books

(Example Book)

"Can you explain, what made Peter blush when ... he was embarrassed?"

Prior Belief Baserate Training Books

(Example Book)

Tra

"How many children got sleepy? Very ... many or very few?"

Control Condition Books

(Example Book)

"What made all of the checkers fall off of the table?"

Figure 2: Percentage of psychosomatic explanations generated by three-and-a-half-year-olds in each of the four conditions.

