When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships

Christopher G. Lucas[1], Sophie Bridgers[2], Thomas L. Griffiths[2], Alison Gopnik[2]

[1]School of Informatics, University of Edinburgh

[2]Department of Psychology, University of California, Berkeley

Word count: NNNN

Keywords: causality, cognitive development, Bayesian models

Address for correspondence:

3.31 Informatics Forum

10 Crichton Street

Edinburgh, EH8 9AB

United Kingdom

E-mail: c.lucas@ed.ac.uk **Phone:** +44 131 651 3260

Abstract

Children learn causal relationships quickly and make far-reaching causal inferences from what they observe. Acquiring abstract causal principles that allow generalization across different causal relationships could support these abilities. We examine children's ability to acquire abstract knowledge about the forms of causal relationships and show that in some cases they learn better than adults. Adults and 4- and 5-year-old children saw events suggesting that a causal relationship took one of two different forms, and their generalization to a new set of objects was then tested. One form was a more typical disjunctive relationship; the other was a more unusual conjunctive relationship. Participants were asked to both judge the causal efficacy of the objects and to design actions to generate or prevent an effect. Our results show that children can learn the abstract properties of causal relationships using only a handful of events. Moreover, children were more likely than adults to generalize the unusual conjunctive relationship, suggesting that they are less biased by prior assumptions and pay more attention to current evidence. These results are consistent with the predictions of a hierarchical Bayesian model.

When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships

Introduction

In everyday life we reason about abstract and general causal principles as well as more concrete and specific causal relationships. For example, I not only know that I have to both put in the plug and push a button on my microwave to make it go, I know more generally that activating appliances has a characteristic causal structure. You both have to ensure that there is power coming to the appliance and that the on/off mechanism is set to "on"; neither cause is sufficient by itself. This sort of abstract principle has been referred to as an *overhypothesis* (Goodman, 1955; Kemp et al., 2007), that is, a hypothesis about the kinds of hypotheses that are likely to be true. Overhypotheses can shape subsequent inferences. If my new digital speaker fails to play, I'll know to check both the power cord and the "play" button on my computer. These abstract principles thus constrain our hypotheses about specific causal relationships and help us learn more effectively (Kemp, Perfors & Tenenbaum, 2007). They play a particularly important role in intuitive theories of biology, physics and psychology as "framework principles" (Wellman & Gelman, 1992; Gopnik & Wellman, 2012). So, where do these principles come from?

Recent work demonstrates that young children are remarkably skilled at learning specific causal relationships (e.g., Gopnik et al., 2004; Kirkham & Sobel, 2007; Kushnir & Gopnik, 2007; Schulz et al., 2007; Gweon & Schulz, 2012). For example, they can infer which blocks will activate a machine based on the contingencies between the blocks and the machine's activation.

But can children also learn more abstract causal principles, and use those principles to shape their subsequent inferences? There is one experiment showing that 4-year-olds can learn abstract causal categories of objects from data (Schulz et al., 2008) and one showing that they can learn abstract psychological categories (Seiver et al, 2013). There is also new evidence that in looking-time experiments, even infants can learn overhypotheses about properties of sets of objects (Dewar & Xu, 2010). There have also been studies examining the development of deductive reasoning and logical rules in children (e.g., Dias & Harris, 1991; Markovits & Vachon, 1990). But there have been no studies examining whether children can learn abstract principles about the logical form of causal relationships, or comparing children's and adults' abilities to do so. In this paper, we show that 4- and 5-year-old children can learn such principles, and can use them to design effective actions. In some circumstances, children learn these abstract causal principles more easily than adults do.

We contrast two abstract causal principles (overhypotheses) about the forms that relationships take in a causal system. One is that relationships have a disjunctive form, in which each cause has an independent probability of bringing about an effect. This form is pervasive in the literature on adult causal inference (e.g., Cheng, 1997; Griffiths & Tenenbaum, 2005). For example, a burglar alarm may be triggered by an intruder or the wind, and a fever may result from a virus or a bacterium. The other overhypothesis is that causal relationships have a conjunctive form in which individual causes are unable to produce an effect, but multiple causes in conjunction can do so. For example, a microwave turns on when both the plug is connected and a button is pressed, but not if either of these causes occurs by itself; likewise, a heart attack may only result if a person has both high blood cholesterol and a particular genetic susceptibility.

Knowing when a machine or a disease has a conjunctive form or a disjunctive form helps us make the right inferences when we want to use the machine or cure patients.

Lucas and Griffiths (2010) showed that adults can learn these overhypotheses about the forms of causal relationships and explained this process in terms of a hierarchical Bayesian model. In a hierarchical Bayesian model, the prior probability of an abstract causal principle is combined with observed data via Bayes' rule. This determines the posterior probability of the principle. The process is hierarchical: evidence can inform both a lower-level hypothesis, such as one about which events are causes of an effect, and an overhypothesis that constrains or leads to that lower-level hypothesis, such as one about how likely causal events are in general, and what kinds of causal relationships apply in a domain. If young children can also learn and then exploit causal overhypotheses, this might help explain the swiftness and generality of early causal learning.

We can also ask whether there are developmental differences between children and adults. Adults appear to be biased towards expecting disjunctive relationships and learn these relationships more easily than conjunctive relationships (Lucas & Griffiths, 2010), a pattern that is consistent with the prevalence of disjunctive relationships in the literature in general (Lu et al., 2008; Griffiths & Tenenbaum, 2005; Cheng, 1997).

Intuitively, we might expect that children would find it more difficult to learn overhypotheses than adults, particularly unusual overhypotheses. After all, dating back to Piaget and Vygotsky, researchers have often assumed that children move from more concrete to more general knowledge. Moreover, adults have both more knowledge and more developed information-processing capacities than children.

However, many developmentalists have recently argued for a Bayesian approach to cognitive development and particularly causal learning (e.g., Gopnik & Schulz, 2007; Gopnik & Wellman, 2012; Tenenbaum et al., 2011; Xu & Kushnir, 2013). The Bayesian approach suggests an alternative and somewhat counterintuitive developmental hypothesis. According to the Bayesian view, learning a new hypothesis involves combining the prior probability of that hypothesis with the observed data. Since children have less experience than adults, their "priors" will be different. In particular, they might be less biased towards hypotheses that are consistent with prior experience and more likely to accept hypotheses – including overhypotheses – that are consistent with new evidence. So children might actually be better at learning an unusual abstract causal principle than adults. In particular, a Bayesian approach suggests that children might find it easier to learn that causal relationships take a conjunctive form.

Differences in the prior expectations of adults and children could take different forms. Adults and children might simply assign high prior probabilities to different hypotheses so that adults have a strong a priori commitment to one kind of relationship, and children to another. Alternatively, adults and children might just differ in the strengths of their commitments, with children holding more diffuse beliefs. This latter possibility is consistent with the difference between a "low temperature" and a "high temperature" system, to borrow an analogy from statistical physics: the adults have congealed in their beliefs and are hard to shift, while the children are more fluid and consequently more willing to entertain new ideas.

If children are more flexible when learning about causal relationships, does this make them better learners than adults? The answer to this question depends on how we define learning. If we say that better learners are people who make correct inferences more often when they are faced with common and familiar situations, then the advantage goes to adults. On the other hand, if we focus on how quickly learners assimilate new information and update their beliefs, and come to understand novel situations, then flexibility – whether it is due less-entrenched ideas about the structure of the world or an exploratory approach to updating beliefs – is a marker of better learning.

There is some reason to believe that children show this sort of superior flexibility in other domains. In cognitive neuroscience, researchers have suggested that young brains, with less top-down control, may be more flexible and plastic than older brains (Thompson-Schill, Ramscar & Chrysikou, 2009). Moreover, young children are able to learn a wider variety of language sounds more easily than adults (Kuhl, 2004), are better than adults at discriminating between faces of non-human primates (Pascalis et al., 2004), and are more likely to look beyond the conventional uses of tools in order to solve problems (German & Defeyter, 2000). However, we do not know whether an analogous effect applies to children's causal learning and their development of intuitive theories.

We examine this developmental hypothesis through head-to-head comparison of children and adults in a causal learning task. Specifically, we explore how children and adults learn that causal relationships follow a conjunctive or a disjunctive form. We examine whether children can form appropriate abstract generalizations, whether they use these abstract principles to shape more specific causal hypotheses and, finally, whether they are more willing to make these generalizations than adults.

Experiment 1: Learning the forms of causal relationships

Young children often have difficulty explicitly articulating causal hypotheses, so we designed a modified version of the experiment in Lucas and Griffiths (2010) that only required yes/no judgments. The experiment had two phases. First, in the training phase, children saw a set of events involving prospective causes ("blickets") and an effect (activation of a "blicketness machine"). One set of events indicated that the machine worked disjunctively – each object individually did or did not activate the machine. The other set of events indicated that the machine worked conjunctively – a combination of two specific objects activated the machine but the individual objects did not.

Next, in the test phase, the participants all saw the same new set of events involving new objects. The test events were ambiguous; they could be consistent with either a disjunctive or conjunctive over-hypothesis. If the participants had inferred the overhypothesis consistent with the evidence observed in the training phase, they should use that hypothesis to interpret the ambiguous test events.

Figure 1 describes the specific events that participants saw in each condition.

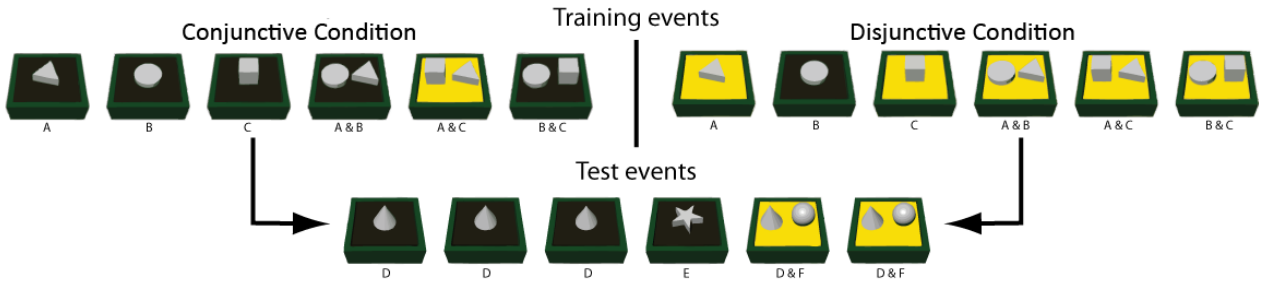

Figure 1. Evidence presented to participants in Experiment 1. In the training phase, participants saw one of the sets of training events, depending on whether they were in the *conjunctive* or *disjunctive* condition. The training events were followed by a set of test events, which all participants saw. Events are given as a set of prospective causes and the presence or absence of an effect. The bright-paneled machines represent events in which the effect occurs and the dark-paneled machines represent events in which the effect does not occur.

In both the training phase and the test, participants should infer that A and C are both blickets and that B is not. However, they should infer different over-hypotheses about the form of the causal relationship, concluding that the machine works conjunctively in the first case and disjunctively in the second. In the subsequent test phase, these overhypotheses should lead to different judgments about which of the new objects (D, E, and F in Figure 1) are blickets.

If children believe the disjunctive overhypothesis and think that a single blicket suffices to activate the machine, they should also believe that object F is likely to be a blicket. However, they should believe that objects D and E are not blickets, since they did not activate the machine by themselves. If, in contrast, they believe the conjunctive hypothesis, they should behave differently. They should believe that F is a blicket, just as they did in the disjunctive case. But they should also believe that D is a blicket since it activated the machine in conjunction with F. They should be uncertain about E, since it did not activate the machine by itself and they have not seen it combined with another object. We will refer to D as the "conjunctively active object", E as the "uncertain object", and F as the "unambiguous object".

There are several developmental possibilities. The training evidence might have no effect on children's test-phase judgments. In that case, we might infer that the ability to form causal overhypotheses is itself a consequence of late-childhood learning or development. Children, like the adults in earlier experiments, might preferentially infer the disjunctive principle, even when the evidence supports the conjunctive principle, and so might tend to call only F a blicket in both cases. In that case, we might infer that a disjunctive bias is in place early in development. Finally, children, unlike adults, might take the training evidence into account equally in both conditions. They might call only F a blicket in the disjunctive condition but call objects F, D and E blickets in the *conjunctive* condition. In that case, we might conclude that the strong adult bias towards disjunctive relationships is due to learning. The adults' experience has led them to assign a higher prior probability to disjunctive overhypotheses.

Participants

Children. Thirty-two children were recruited from university-affiliated preschools, divided evenly between the *conjunctive* and *disjunctive* conditions. Children in the *conjunctive* and *disjunctive* conditions had mean ages of 4.46 (4.02-4.85; SD=0.27) and 4.61 (4.13-4.99; SD=0.31) years, respectively.

Adults. UC Berkeley undergraduates received course credit for participating during lectures of an introductory psychology course. There were 88 participants in the *conjunctive* condition and 55 in the *disjunctive* condition. Five participants in the *conjunctive* condition were excluded for declining to answer questions.

Methods

Children. Each child sat at a table facing the experimenter, who brought out three gray ceramic objects, each with a different shape, as well as a green box with a translucent panel on top, describing the box as “my blicketness machine”.

At the beginning of the experiment, children were prompted to help the experimenter name the objects using their shapes, e.g., “triangle”. They were then told that the goal of the game was to figure out which of the objects were blickets, and that blickets cannot be distinguished from non-blickets by their appearance. They were also told that blickets have blicketness inside them. This was designed to encourage the conjunctive interpretation – two blickets might be necessary to accumulate a critical amount of blicketness needed to activate the machine. No other information was provided about the blickets or the machine.

The children then observed a set of training events in which the experimenter placed objects alone or in pairs on the machine. In some cases the machine activated by lighting up and playing music. These events corresponded to either the *disjunctive* condition or the *conjunctive* condition training given in Figure 1. After the children saw these events, the experimenter asked whether or not each object was a blicket. Next, the experimenter brought out three objects that the children had not seen before. After the children named the new objects by their shapes, the experimenter demonstrated the test events listed in Figure 1 and asked whether or not each of these new objects was a blicket.

The experiment was repeated a second time for each child, using the same patterns of evidence, but with a distinct set of objects that varied in their colors rather than shapes. Both the identities of the individual objects that activated or did not activate the machine and the order of the sets were counterbalanced.

Adults. The adults were tested in groups using a procedure that was identical except that the adults were not asked to name the objects, and they recorded their judgments on sheets of paper rather than responding verbally.

Results

Children. If children are (1) learning about the form of the relationship between blickets and the machine's activation, and (2) transferring that abstract knowledge to make inferences about novel ambiguous events, then this would lead to the following behaviors. In the *disjunctive* condition, they should say that F is a blicket more often than they say that D and E are blickets.

In the *conjunctive* condition, they should be likely to say that F and D are blickets. They should also say that both D and F are blickets more often than they say that E is a blicket.

Comparing the two conditions, children should say that F is a blicket equally often in both conditions, but that D is a blicket more often in the *conjunctive* than in the *disjunctive* condition. Children might also say that E is a blicket more often in the *conjunctive* than *disjunctive* condition, since E is definitely not a blicket in the disjunctive case, but whether or not it is in the conjunctive case is uncertain[1].

These predictions were largely borne out. Children in the *disjunctive* condition chose the unambiguous object F as a blicket significantly more often (M=1.5 of 2, SD=.63) than object D (M=.38 of 2, SD=.62; $p < .001$, McNemar's exact test, one-tailed) or E (M=.50 of 2, SD=.73; $p = .003$, McNemar's exact test, one-tailed). In the *conjunctive* condition, they did not choose object F (1.63 of 2, SD=.62) more often than they chose object D (M=1.31 of 2, SD=.79; $p = .51$, McNemar's exact test). They chose object F more often than they chose the uncertain object E (M=.75 of 2, SD=.77; $p = .004$, McNemar's exact test, one-tailed) and also chose object D more often than object E ($p = .035$, McNemar's exact test, one-tailed).

Children also judged object D to be a blicket more often in the *conjunctive* condition than in the *disjunctive* condition ($p = .001$, one-tailed permutation test), though they were equally

---

[1] We used one-tailed tests to evaluate these specific predictions, and 2-tailed tests to assess the significance of findings when the direction of effect was not predicted.

likely to say that F was a blicket in both conditions (81 percent and 75 percent of judgments in the *conjunctive* and *disjunctive* conditions, respectively, (p = .78, two-tailed permutation test)). Children also tended to say that E was a blicket more often in the *conjunctive* than *disjunctive* condition, but not at a significant level (p = .24, one-tailed permutation test). See Figure 2, top row, for a summary of ratings in the four conditions.

Adults. Adults showed the same pattern as children in the *disjunctive* condition. They said that F (M=1.95 of 2, SD=.23) was more likely to be a blicket than D (M=.13 of 2, SD=.34; $p < .001$, McNemar's exact test, one-tailed) or E (M=.28 of 2, SD=.63; $p < .001$, McNemar's exact test, one-tailed). In the *conjunctive* condition they behaved differently, however, saying that F (M=1.52 of 2, SD=.70) was significantly more likely to be a blicket than D was (M=.47 of 2, SD=.77; $p < .001$, McNemar's exact test), rather than saying that D and F were both likely to be blickets. In fact, they said that D was not a blicket significantly more often than half the time ($p < 0.001$, $t(82) = 6.269$).

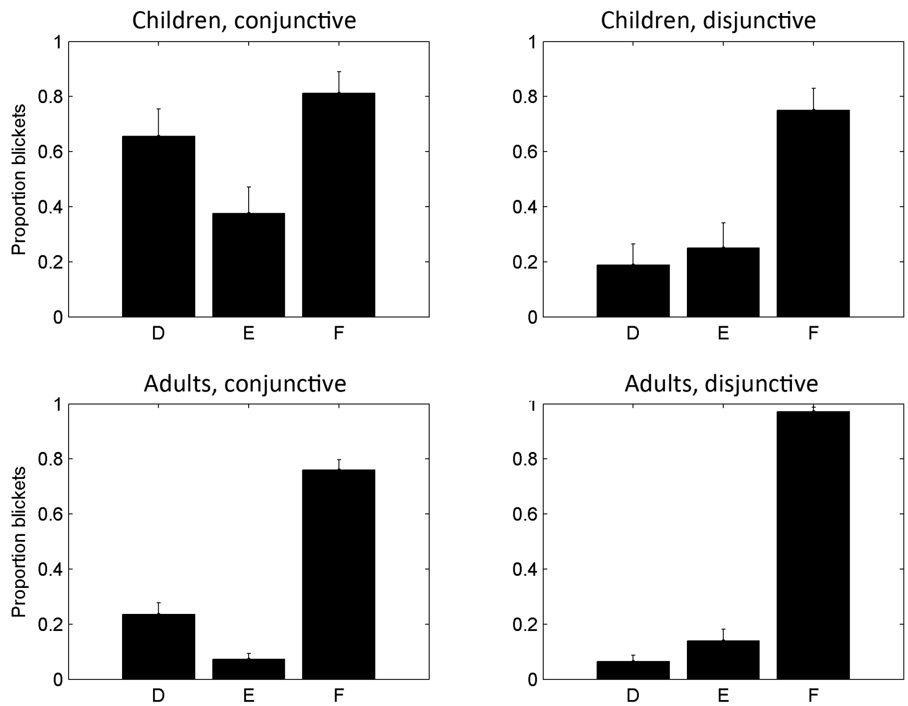

Figure 2. Proportions of the objects D (the conjunctively active object), E (the uncertain object), and F (the unambiguous object) that were judged to be blickets in Experiment 1. Proportions for children are in the top row and judgments for adults are in the bottom row, for the *conjunctive* (left column) and *disjunctive* (right column) conditions. Error bars represent standard errors of the mean.

However, adults were also somewhat more likely to judge object D, the conjunctively active object, to be a blicket in the *conjunctive* condition than in the *disjunctive* condition (p = .0031, two-tailed permutation test). This finding is consistent with the results in Lucas and Griffiths (2010), and suggests that adults were also somewhat sensitive to the training evidence, though less so than the children See Figure 2, bottom row, for a summary of adults' judgments for the test objects.

Differences between children and adults. There were no significant differences between children and adults in the *disjunctive* condition. However, in the *conjunctive* condition, the children (first-repetition M=.63, SD=.50) judged the conjunctively active object D, to be a blicket more frequently than adults (first-repetition M=.22, SD=.41; p = .0019, Fisher's exact test). If there is a conjunctive relationship, then object D, which activates the machine in combination with F, is likely to be a blicket.

Children's ratings (first-repetition M=.44, SD=.51) were also significantly higher than adults' (first-repetition M=.05, SD=.22) for the uncertain object E (p < .001, two-tailed permutation test). If there is a conjunctive relationship, then the event where object E fails to activate the machine is uninformative, so the child's judgment about whether object E is a blicket should reflect how likely blocks are to be blickets. Furthermore, if there is a conjunctive relationship then 4 of the 5 objects are blickets – a fairly high base rate. This would lead the children to judge that object E was somewhat likely to be a blicket.

This pattern of results thus suggests that children were more likely to infer a conjunctive relationship than adults. The children showed a stronger discrimination between the two conditions than the adults did.

Note that the children's performance on the uncertain E block also makes it unlikely that children in the conjunctive condition are simply confused and therefore responding with a "yes" bias. Recall that children were less likely to say that object E was a blicket than that objects D or F were. If they had simply been responding with a "yes" bias they should have said that all the blocks were blickets.

Discussion

The children in Experiment 1 took the training data into account equally in both conditions, but the adults did not. Both children and adults behaved the same way in the disjunctive condition. They said that F was a blicket and D and E were not. However, in the *conjunctive* condition, children responded as if D was also a blicket, and E might be a blicket. In contrast, adults' judgments were only weakly influenced by the *conjunctive* training data; they continued to say that F was much more likely to be a blicket than D or E. This pattern supports the hypotheses that (1) children are able to learn that a causal relationship is conjunctive, and (2) do so more readily than adults.

We can also exclude some alternative explanations for these results. One alternative explanation is that children might be more likely than adults to judge any object to be a blicket, across both conditions. In fact, however, adults were more likely than children to call object *F* a blicket in the *disjunctive* condition, and nearly as likely in the *conjunctive* condition (75 percent of the objects versus 81 percent). Children only showed an increase for the D and E blocks.

A second alternative is that the children were confused by the training data in the *conjunctive* condition, and responded to the novel objects by guessing randomly. This explanation can be ruled out by noting that children judged objects D and F to be blickets more often than chance would predict ($t(15)=3.529$, p = .0030)[2]. A third alternative is that the children were confused by the training data in the conjunctive condition and so responded with a "yes" bias. However, the fact that children chose E as a blicket less often than D or F weighs against this interpretation.

Some alternative explanations are still possible. First, it is possible that children in the *conjunctive* condition were simply confused, and so resorted to using associations between objects and the effect in the test condition. In the test condition, both D and F were associated with activation twice, while E was never associated with activation, and this may have led children to prefer to say that D and F were blickets and E was not.

Another possible explanation for our results is that the questions are not really distinguishing between the children's and adults' causal beliefs, but rather tests their use of terms like "blicket". Our results cannot be explained in terms of a simple "yes" bias and we used the "blicketness" terminology to encourage participants to consider the conjunctive possibility. However, children might still be more willing than adults to call an object a blicket, even though adults and children make a similar inference about the causal relationship itself. In particular, adults might appreciate that the blocks are conjunctive causes, but still be reluctant to call such causes "blickets" because they are sensitive to linguistic nuances that children ignore.

If this explanation is correct, adults and children should behave similarly when they are asked to make the machine go. However, if the children really had inferred different causal

---

[2] Individual comparisons for D and F yielded trends, but an aggregate analysis suffices to discount the explanation that children were guessing randomly in general.

structures in the two conditions they should use single blocks to make the machine go in the *disjunctive* condition but should try combinations of blocks in the *conjunctive* one. Similarly, if the adults really had preferentially inferred the disjunctive structure they should use single blocks to make the machine go in both conditions. We conducted Experiment 2 to address these possibilities.

Experiment 2: Interventions and baselines

Experiment 2 made several changes to the design of Experiment 1 in order to rule out alternative hypotheses.

In Experiment 1, we asked participants if objects were blickets. Children and adults might treat this question differently. However, if someone genuinely believes that an event X causes an event Y, then she should produce X to bring about Y. Therefore, in Experiment 2 we asked participants which objects they would use to activate the machine.

As we noted above, one possibility is that children in the *conjunctive* condition had trouble interpreting the training events and fell back on an associative strategy. In the original test events, D- D- D- E- DF+ DF+, the unambiguous F object was associated with two activations of the machine and no non-activations, the conjunctively active object D was associated with three non-activations and two activations, and the uncertain object E was associated with one non-activation and zero activations. A strategy of relying on these associations might explain the difference between D, F and E judgments. The relatively high rate of blicket judgments might be puzzling given that children called object E, which had no association with the machine, a blicket almost half the time, but this could be attributed to a general bias toward calling objects blickets. We took several steps in Experiment 2 to address these alternative explanations.

First, we provided children and adults with evidence that blickets are rare in comparison to non-blicket objects at the outset of the experiment. This was intended to attenuate any yes-bias that might be influencing the judgments of the children.

Second, we modified the test events so that the machine had the same probability of activating in the presence of objects D and E. The test events for Experiment 2 were D- D- D- E- DF+ DEF+ DF+, so that D was associated with three activations and three non-activations, and E was associated with one activation and one non-activation. If children simply relied on associations there should be no difference between judgments about D and E. Moreover, this modification allowed us to test whether our results would replicate with a different set of test events.

Third, we introduced a new object, G, which we will call the "novel object" which did not participate in any events. Children and adults were asked to judge whether G was a blicket. This allowed us to estimate empirically how likely children and adults were to judge that a novel object was a blicket when no evidence was available. In turn, this also allowed us to ensure that responses to D and F were not somehow due to participants' beliefs about the baseline probability that object were blickets.

Fourth, we added a new baseline condition, where the training events were omitted. This condition provides us with a clearer picture of how the conjunctive and disjunctive training events shape the inferences that participants make about the test objects. This also provides another control – it allows us see if the conjunctive pattern in children is due to the conjunctive training or is simply a default pattern, which might emerge if the children simply ignored the training trials.

Finally, we simplified the procedure by giving participants one test trial instead of two. We provided each participant with two repetitions of the training phase before a single test phase.

Participants

Children. Seventy-four children were recruited from university-affiliated preschools and local children's museums, and were divided between the *conjunctive* (n=25), *disjunctive* (n=25), and *baseline* (n=24) conditions. Children in the *conjunctive*, *disjunctive,* and *baseline* conditions had mean ages of 4.82 (4.04-6.00; SD=.54), 4.80 (3.72-5.90; SD=.65), and 5.04 (3.90-5.84; SD=.62) years, respectively. A total of nine participants were replaced due to experimenter error, including one in the *conjunctive* condition, two in the *disjunctive* condition, and six in the *baseline* condition.

Adults. UC Berkeley undergraduates received course credit for participating in groups of up to five students. There were 28, 28, and 26 participants in the *conjunctive*, *disjunctive*, and *baseline* conditions, respectively. Five participants were replaced due to experimenter error in the *baseline* condition and no participants were replaced in the *conjunctive* condition nor in the *disjunctive* condition.

Methods

Children. The methods resembled those from Experiment 1, with the following changes. We added a base rate manipulation to the beginning of the experiment, in which children saw evidence that blickets were rare. The experimenter told participants that only a few of the objects were blickets and that most of them were not. To further illustrate this, the experimenter produced two different buckets of objects, one labeled “Blickets” containing one object and a second labeled “Not Blickets” containing four objects. The experimenter asked each participant to help count the number of objects in each bucket before noting that there were many more non-blickets than blickets.

In all three conditions, before each set of the training or test events, the experimenter drew the test or training objects from a bucket of unsorted objects, apparently at random, and prompted participants to name the objects using their shapes. The identities of the individual objects used in the training and test events were counterbalanced.

In the *conjunctive* and *disjunctive* conditions, after the base rate manipulation, children observed two sets of training events that corresponded to either the *conjunctive* condition or *disjunctive* condition training shown in Figure 1. After demonstrating each set of events, the experimenter asked whether or not each object involved in the set was a blicket. Next, the experimenter discovered an object (G) that she had “forgotten” and, without putting the object on the machine, asked children whether or not they thought this object was a blicket. Afterward, the experimenter drew the last three objects from the bucket and asked the children to name them. The experimenter then demonstrated the new test events: D- D- D- E- DF+ DEF+ DF+. After demonstrating these events, the experimenter asked whether or not each object was a blicket. Lastly, children were asked the intervention question, "Which of these should we use to make the machine turn on?" which prompted children to say which of the test objects they would use to make the machine activate.

In the *baseline* condition, children were given two sets of test events without any training events. After the base rate manipulation described above, the experimenter asked children about the “forgotten” object G and then demonstrated the test events. After each set, children were asked whether the objects were blickets, followed by the intervention question, exactly as in the other two conditions.

Adults. The adults were tested in groups using a procedure that was identical except that the adults were not asked to name the objects and recorded their judgments on sheets of paper rather than responding verbally.

Results

Predictions. As in the previous experiment, we predicted that children in the *disjunctive* condition should say that F was a blicket and D and E were not. In the *conjunctive* condition they should say that F and D were blickets and be uncertain about E. In terms of interventions, if children believe that the machine operates on a disjunctive principle, and thus infer that F is a blicket and D is not, they should tend to place F on the machine rather than D or E. Moreover, they should be more likely to place single objects on the machine than multiple objects. In contrast, if they believe the machine operates on a conjunctive principle they should place both D and F on the machine, and might also experiment with the uncertain object E. They should also put multiple rather than single objects on the machine.

Children. As predicted, Children in the *disjunctive* condition were significantly more likely to call the unambiguous object F a blicket (M=.80, SD=.41) than either the conjunctively active D (M=.32, SD=.48; $p = .001$, McNemar's exact test, one-tailed) or the uncertain E (M=.28, SD=.46; $p < .001$, McNemar's exact test, one-tailed). They were also equally likely to call D and E blickets ($p = 1.0$, McNemar's exact test).

In contrast, in the *conjunctive* condition children were equally likely to call F and D blickets, and they were more likely to call both objects F (M=.88, SD=.33) and D (M=.92, SD=.28) blickets than object E (M=.68, SD=.48; $p = .031$ and $p = .016$, respectively, McNemar's exact test, one-tailed). They also called both D and F objects blickets at greater than chance levels ($p < .001$ in both cases, one-tailed binomial tests). This is consistent with our predictions, and contrary to the predictions of an associative learning model.

As seen in Figure 3, children were more likely to call the conjunctively active object D a blicket in the *conjunctive* condition than they were in the *disjunctive* condition (one-tailed Fisher's exact test, $p < .001$), and the *baseline* condition (M=.42, SD=.50; one-tailed Fisher's exact test, $p < .001$). There were no significant differences in their judgments of D, E (baseline M=.33, SD=.48), or F (baseline M=.75, SD=.44) between the baseline and disjunctive conditions.

Children's choices of interventions revealed the same patterns as their blicket versus not-blicket judgments. Children in the *conjunctive* condition had a strong tendency to choose interventions that included the conjunctively active object D, doing so significantly more than half of the time (20 of 25; one-tailed binomial test, $p = .002$). They also chose interventions involving multiple objects (21 of 25; one-tailed binominal test; $p < .001$), indicating that they believed that a conjunctive relationship was at work. The rate of choosing interventions involving D was significantly lower in the *disjunctive* condition (Fisher's exact test, one-tailed, $p < .001$), where only 3 of 25 children chose interventions including object D. Similarly, children in the *disjunctive* condition never chose multiple objects. See Figure 3 for a summary of the children's judgments.

Adults. Adults in the *disjunctive* condition behaved much like the children, choosing the F object (M=.82, SD=.39) significantly more than D (M=.11, SD=.31) or E (M=.11, SD=.31; $p <$ .001 in both cases, McNemar's exact test). Adults in the *conjunctive* condition, however, behaved differently. They were not significantly more likely to describe object D (M=.25, SD=.44) as a blicket than object E (M=.11, SD=.31; McNemar's exact test, $p = .22$), suggesting that they were less strongly influenced by the *conjunctive* training than children. Adult participants were more likely to choose F (M=.71, SD=.46) than D ($p = .004$, McNemar's exact test), in the *conjunctive* condition. Adults were somewhat more likely to call object D a blicket in the *conjunctive* condition than they were in the *disjunctive* condition (Fisher's exact test, $p = .29$), and the *baseline* condition (Fisher's exact test, $p = .14$), but this difference did not reach significance. See Figure 3 for a summary of the adults' judgments.

As with the children, adults' interventions were consistent with their blicket and not-blicket judgments, indicating that the difference between the two groups was due to a difference in their causal beliefs rather than a difference in how they interpreted the word "blicket". Just as adults tended to say that only object F was a blicket, by far their most frequent response in all three conditions was to activate the machine with F alone. In the *conjunctive* condition, adults tended not to include object D in their interventions (9 D, DE, DF or DEF responses of 28) and their interventions tended to involve single objects (18 of 28).

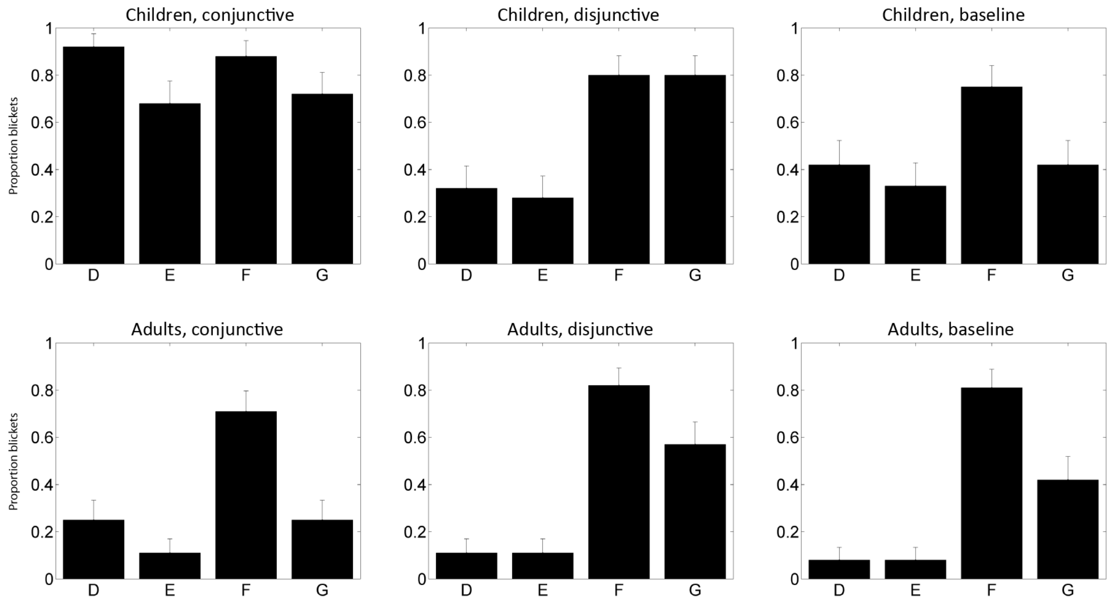

Figure 3. Proportions of objects in Experiment 2 that were judged to be blickets for children (top row) and adults (bottom row) for the *conjunctive* (left column), *disjunctive* (center column) and *baseline* (right column) conditions. Error bars represent standard errors of the mean.

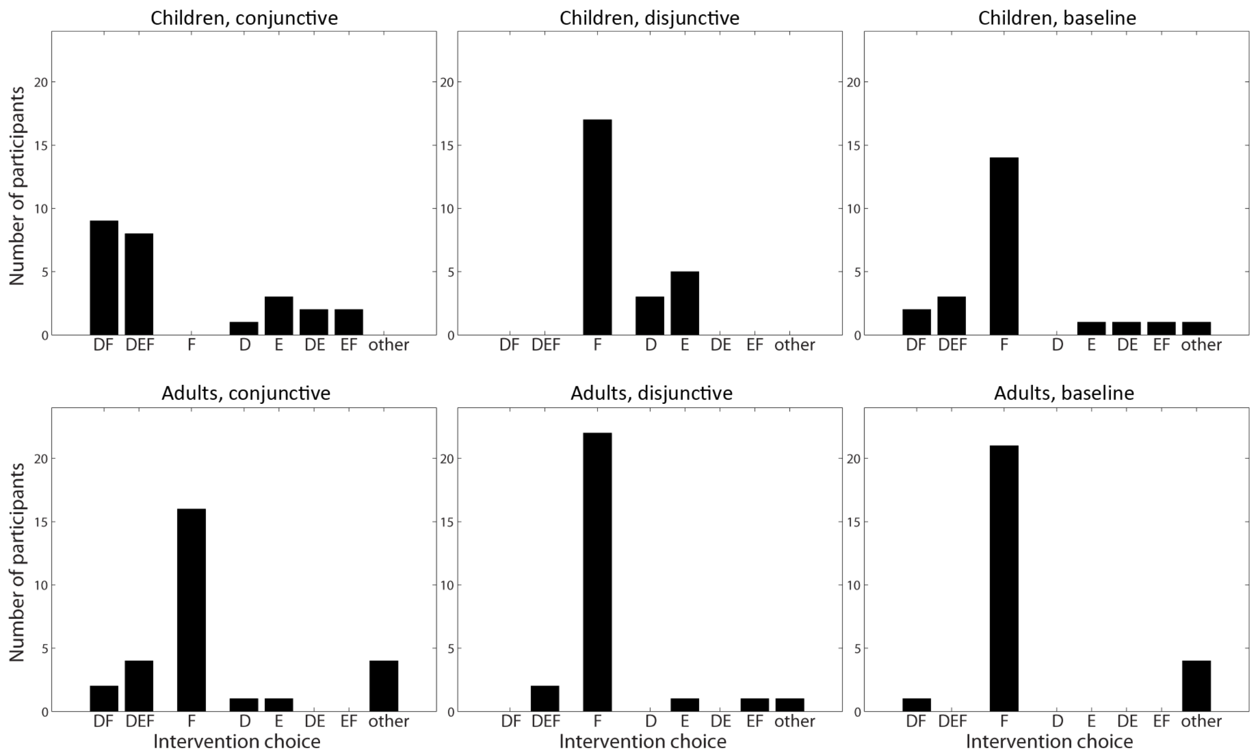

Figure 4. Participants' choices of interventions for Experiment 2, including children (top row) and adults (bottom row) for the *conjunctive* (left column), *disjunctive* (center column) and *baseline* (right column) conditions. Intervention choices are labeled by the objects participants chose, e.g., DEF indicates a participant chose objects D, E, and F. Intervention choices are ordered by the relationship types that they indicate, from left to right: conjunctive (DF and DEF), disjunctive (F), and neither.

Differences between children and adults. As predicted, children called the conjunctive active object D a blicket more frequently than adults did in the *conjunctive* condition (Fisher's exact test, one-tailed, $p < .001$), there was no significant difference in their judgments of F (Fisher's exact test, $p = .18$). In the *disjunctive* condition, there were no significant differences in children's and adults' responses to the D, E, and F objects (p > .05, Fisher's exact test).

In choosing interventions in the *conjunctive* condition, children were more likely to include object D than adults (Fisher's exact test, $p < .001$), and were more likely to choose multiple objects (Fisher's exact test, $p < .001$). This supports the hypothesis that most children are inferring that the underlying relationship is conjunctive, while most adults are not.

Responses to the novel object G. Recall that the participants were asked to judge whether the novel "forgotten" object G was a blicket after the training trials but before the test trials. We added the G object to determine participants' baseline beliefs about an unknown object being a blicket. We predicted that children would say that F was a blicket significantly more than G in both conditions, and that they would say that D was a blicket more often than G in the *conjunctive* but not the *disjunctive* condition.

In the *baseline* condition, with no training trials, both children and adults said that G was a blicket about half the time, and there was no significant difference between them. This suggests that, in spite of the initial manipulation, participants did not actually conclude that blickets would be rare in the test itself. In the *disjunctive* condition, with somewhat more positive evidence that objects were blickets, both children and adults were somewhat more likely to call G a blicket, and there was a trend towards children being more likely to call G a blicket than adults (Fisher’s exact test $p = .088$). In the *conjunctive* condition, children were significantly more likely than adults to say that object G was a blicket (Fisher's exact test, one-tailed, $p < .001$). We address explanations for this pattern in the general discussion.

However, it is also notable that in the *conjunctive* condition, child participants, as predicted, described object D as a blicket more often than the novel object G ($p = .031$, McNemar's exact test, one-tailed). In contrast, in the *disjunctive* condition, they were less likely to call object D a blicket than the novel object G ($p = .001$, McNemar's exact test, one-tailed). Thus, the children’s judgments about object D were not simply the result of a default tendency to call objects blickets at a particular rate, a conclusion that is also supported by our intervention data.

Discussion

The results of Experiment 2 support the hypothesis that children are learning conjunctive relationships where adults are not. The contrast between judgments about objects D and E shows that children in the *conjunctive* condition did not simply rely on the association between object D and the effect. Children and adults also made strikingly different interventions in the *conjunctive* condition, so the difference between children's and adults' inferences is not merely due to different beliefs about how the word "blicket" should be used.

The results of this experiment, particularly the intervention results, provide strong evidence that children respond to the training evidence and are quick to infer both conjunctive and disjunctive over-hypotheses, depending on the data they observe. While children in the *disjunctive* condition made similar inferences to those in the *baseline* condition, this result is consistent with the idea that like adults, children prefer disjunctive hypotheses a priori, but are more flexible and less constrained by their prior beliefs.

Adults, in contrast, showed a disjunctive bias. However, one might wonder about the generality of the adult bias. It may be that our “blicket” domain, involving electrical or mechanical devices, is unusual — perhaps adults have especially strong expectations that artifacts or electrical systems are disjunctive. Another possible issue is that adults' biases were the result of their sensitivity to linguistic nuances that children miss. It could be that by using nouns to describe the blocks, that is, by distinguishing between blickets and objects that are not blickets, as opposed to distinguishing “blicket blocks” and “non-blicket blocks” we inadvertently signaled that the underlying causal relationships were disjunctive. Children, failing to recognize this distinction, would not show the same bias as adults.

If our results for adults are specific to stimuli that are artifacts or electrical devices, then changing our cover story should make adults infer a conjunctive relationship in the *conjunctive* condition. If our results are specific to situations where causes are described by nouns then using adjectives rather than nouns to name the objects, e.g., saying "blicket block" rather than "blicket" should also make adults infer conjunctive relationships in the *conjunctive* condition.

Experiment 3: Domain and language controls for adults

We designed a new experimental condition to test whether the syntactic and semantic details of our cover story were responsible for the adult disjunctive bias. Our first goal was to determine whether the blicket cover story – and electrical devices in general – caused the bias. Our second goal was to determine whether picking out causes using nouns ("blickets") rather than adjectives ("blicket blocks") caused the bias. With those goals in mind, we repeated the *conjunctive* condition of Experiment 2 using a modified cover story. Specifically, we replaced the blocks and the machine with flowers that could potentially make you sneeze. We also changed the language, asking participants to judge which flowers were "tulver flowers" and which were not.

Participants

Carnegie Mellon University undergraduates received course credit for participating in groups of up to five students. There were 27 participants in the single *conjunctive* condition after replacing one group of 4 students due to experimenter error.

Methods

The experimental procedure and materials followed those used for adults in the *conjunctive* condition of Experiment 2, including the base rate manipulation, two training phases and one test phase with the same sequence of events, and classification and intervention questions.

However, the cover story, modeled on Schulz and Gopnik (2004), was biological rather than mechanical. The effect was that a toy bear sneezed, and the prospective causes were differently colored flowers. Whereas in Experiments 1 and 2, participants were told that some objects were blickets and some were not, participants in Experiment 3 were told that some flowers were "tulver flowers" which had "tulverness inside them" and that "tulverness makes bear sneeze." This cover story should be compatible with both a disjunctive account and a conjunctive account in which the accumulated strength of an allergen leads to sneezing. At every point where "blicket" was used in Experiment 1, "tulver flower" was used in Experiment 2. The protocols were otherwise identical except that in the intervention choice question, participants were asked which flowers they should keep away from the bear to prevent him from sneezing, rather than which flowers should be used to cause the effect.

Results and discussion

Participants were much more likely to judge flower F to be a tulver flower than flower D ($p < .001$, McNemar's exact test, one-tailed), with 25 of 27 calling F a tulver flower, and zero calling D a tulver flower. The same pattern held for their choices of interventions, in this case the flowers they would remove to prevent bear from sneezing. Twenty-one of 27 participants said to remove only flower F, one said to remove flower D, and one said to remove flowers E and F. The remaining four did not refer to any of the training or test objects in their answers, but instead said "tulver flowers", "all flowers that mix to make black", or the "[base rate manipulation flower] in the box called 'tulver flowers'".

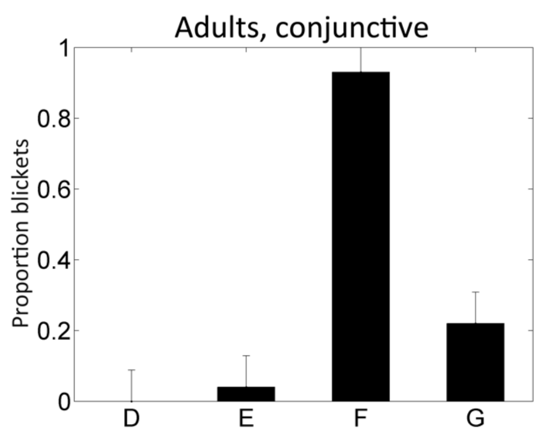

Figure 5. Proportions of objects in Experiment 3 that participants judged to be tulver flowers. Error bars represent standard errors of the mean.

We found that adults were, if anything, less likely to infer a conjunctive relationship when causes were tulver flowers that produced sneezing than when causes were blickets that activated a machine. The disjunctive bias emerged even with different objects, a different causal relationship, and different language.

General Discussion

Our experiments were designed to explore two questions: whether children could learn high-level generalizations about the form of causal relationships at all, and, if they could, how they differed from adults. Our results show that children can learn the forms of causal relationships, and also that they can be more sensitive to evidence than adults are.

This result might seem surprising. After all, adults have more working memory and attentional resources than children and have seen a wider variety of causal relationships. A hierarchical Bayesian account, however, provides a clear explanation for our results. Adult experience is a double-edged sword. We constantly acquire abstract knowledge about the causal structure of the world around us. This provides us with inductive biases that are usually helpful – they let us draw quick and confident conclusions when a new system is consistent with our past experience. However, when we encounter unusual causal systems – like the *conjunctive*-condition blicketness machine and tulver flowers, these same biases can make us reluctant to revise our beliefs. Plausibly, more common causal systems have a disjunctive than a conjunctive structure.

We have focused so far on the qualitative predictions that follow from a hierarchical Bayesian perspective. In the remainder of the paper, we go a step further and compare human judgments to the detailed predictions of a specific hierarchical Bayesian model. This comparison allows us to better understand how the prior beliefs of children and adults might differ. We can also see whether the specific assumptions in our model capture aspects of how children and adults acquire and use abstract causal knowledge. We also consider an alternative formal explanation for our results. Children and adults may differ in the extent to which they favor "exploration" as opposed to "exploitation" in their decision-making.

A hierarchical Bayesian model

In general, Bayesian models begin with a set of assumptions about what events a given hypothesis predicts, and how likely different hypotheses are to be true in the first place. In this case, we start with the same basic assumptions that Lucas and Griffiths (2010) made in predicting adult judgments in similar tasks to ours. The first assumption is that only blickets influence whether or not the machine activates, and that blickets can never decrease the probability that the machine will activate. The second assumption is that there is no variation among blickets – no blicket is more effective than any other blicket. Based on these assumptions, we can express the probability of the machine activating in terms of the number of blickets that are present, n, where the probability of activation increases with n.

The specific form of this relationship is determined by overhypotheses – beliefs that span multiple contexts and determine how data are interpreted. For example, if a learner believes that single blickets deterministically cause the machine to activate, that leads to a function where the probability of activation is 1 for all n greater than zero.

How can we define our space of overhypotheses so they include many possibilities without being too complex? One approach is to use the family of logistic functions, which captures a wide range of relationships, including different types of disjunctive and conjunctive relationships. This family has two parameters, the gain and the bias, with gain corresponding to how reliable the relationship is, and bias corresponding to how many blickets must be present before the machine is likely to activate. An overhypothesis can be expressed in terms of these parameters. The probabilities of different parameter values reflect a learner's a priori beliefs about which relationships are more or less likely.

To test the idea that children are more flexible because they entertain a wider range of possibilities, we explored different distributions of the gain and bias. Children might be more flexible because they have more diffuse expectations than adults. In that case children's judgments should be consistent with models in which the gain and bias have higher variance, while the models that fit adult behavior should have lower variance. In contrast, children and adults might have different but comparably strong commitments. Then, the best priors for children and adults should have similar variances and different means.

There were strong differences between the bias distributions that predicted adults' and children's performance. Adults' judgments were consistent with very strong commitments to one blicket being sufficient to cause the effect (with an expected bias of .2, and a variance of .05), while children's judgments were consistent with the idea that they are amenable to a wider range of relationships (with an expected bias of 1.2, and a variance of 1.9). Both groups' judgments were consistent with diffuse distributions over gains, which was expected given that our experimental manipulations did not focus on biases toward or against reliable relationships. For a sense of the expectations these parameter values encode, see Figure 6, which shows samples for both groups' best-fitting priors.

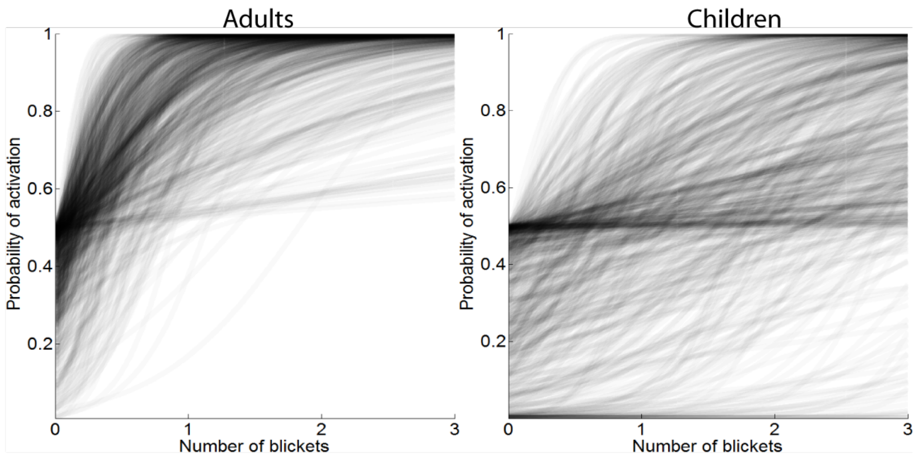

Figure 6. Samples from the priors that best fit adult and child judgments. Each plot shows 600 relationships that have been sampled from each of these priors, revealing the much wider range of likely relationships under the child-fitted priors.

The model must also make assumptions about how likely it is that a new object is a blicket. The simplest such assumption is that new objects have a certain fixed probability, e.g., p(blicket)=.5, which was the approach taken in Lucas and Griffiths (2010). We require a somewhat more sophisticated approach, because we expect that participants learn how common blickets are over the course of the experiment – recall that in Experiment 2, we told participants that one of five novel objects were blickets initially, but they then saw that 4 of 5 were blickets in the training. To accommodate this kind of learning, we expanded the model in Lucas and Griffiths to include overhypotheses about how common blickets are, in the form of distributions over p(blicket). By exploring different distributions over p(blicket), we were able to better understand how well both groups' judgments could be explained by a "yes bias", or a tendency to expect new objects to be blickets.

We found no evidence for such a bias – instead, the priors that were most consistent with children's and adults' responses favored rates of blickets near .5. The main difference was that adults were more likely to make judgments consistent with expecting more extreme rates, high and low, of blicket occurrence.

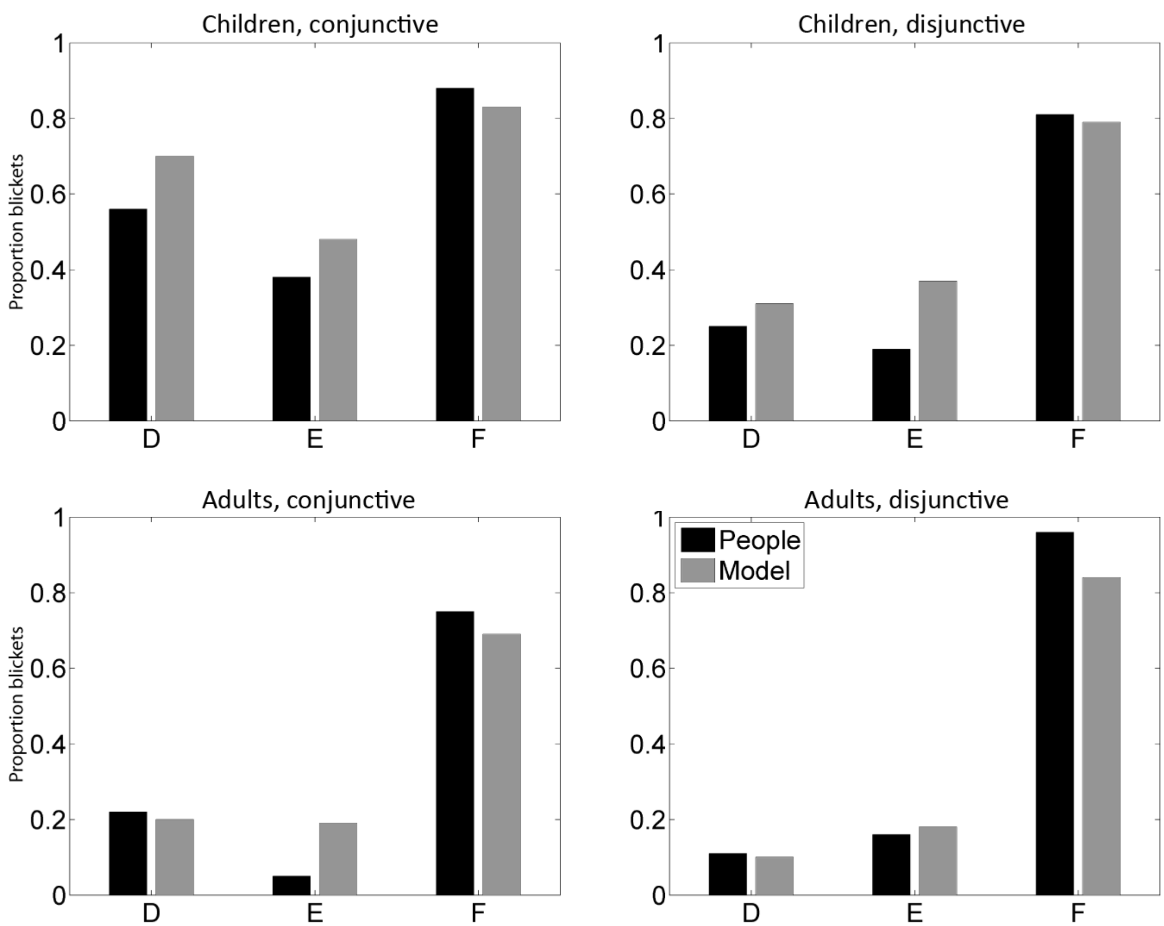

Figure 7. Proportions of objects in Experiment 1 that were judged to be blickets for children (top row) and adults (bottom row) for the *conjunctive* (left column), and *disjunctive* (right column) conditions, along with predictions from our model.

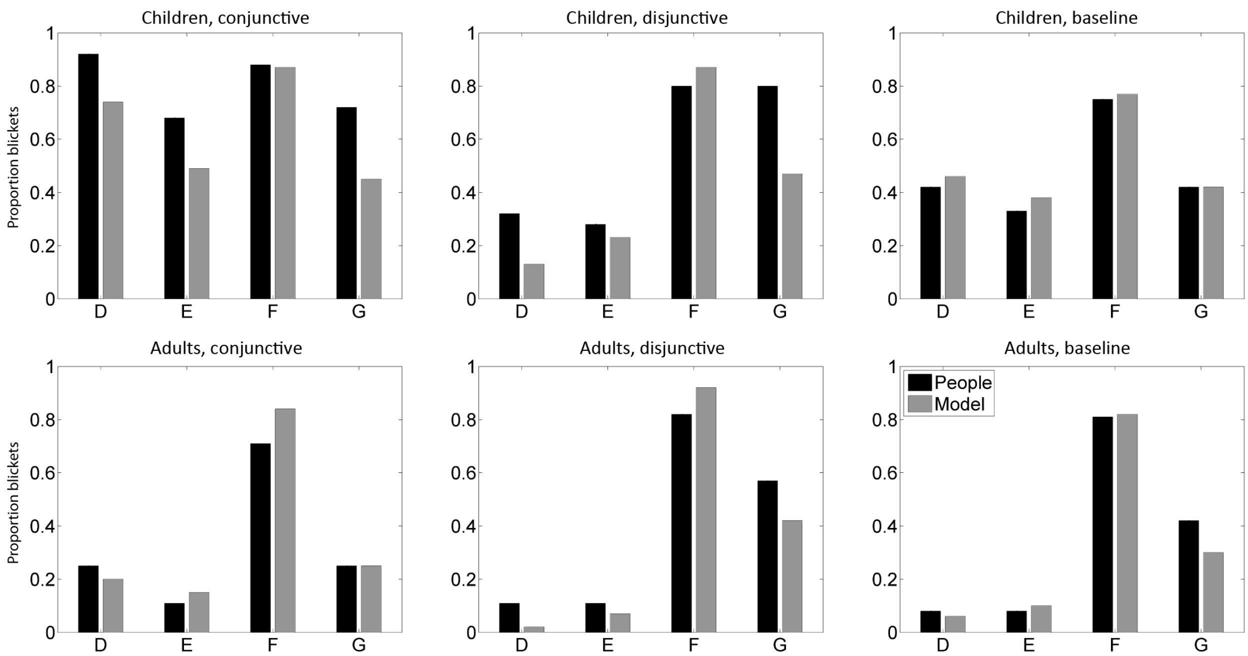

Figure 8. Proportions of objects in Experiment 2 that were judged to be blickets for children (top row) and adults (bottom row) for the *conjunctive* (left column), *disjunctive* (center column) and *baseline* (right column) conditions, along with predictions from our model.

Figures 7-9 show model predictions and results for individual experiments, which are aggregated in Figure 10. They reveal that the priors we have described capture the overall pattern of judgments for both children and adults, with one exception: children tended to expect the novel objects (G) in the *conjunctive* and *disjunctive* conditions of Experiment 2 (but not the *baseline* condition) to be blickets, whereas adults did not. One explanation for this difference is that children discounted the base-rate manipulation. This may have been because those events were qualitatively different than the main training and test events. It is also possible that the children tended to forget the results of base-rate manipulation by the time they were asked about object G, due to their more limited working memory. See Appendix A for details of our modeling results.

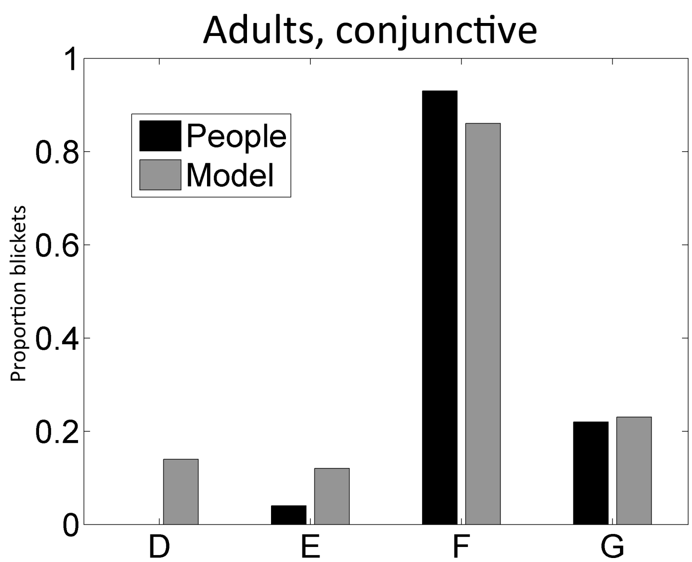

Figure 9. Proportions of objects in Experiment 3 that participants judged to be tulver flowers, along with predictions of our model.

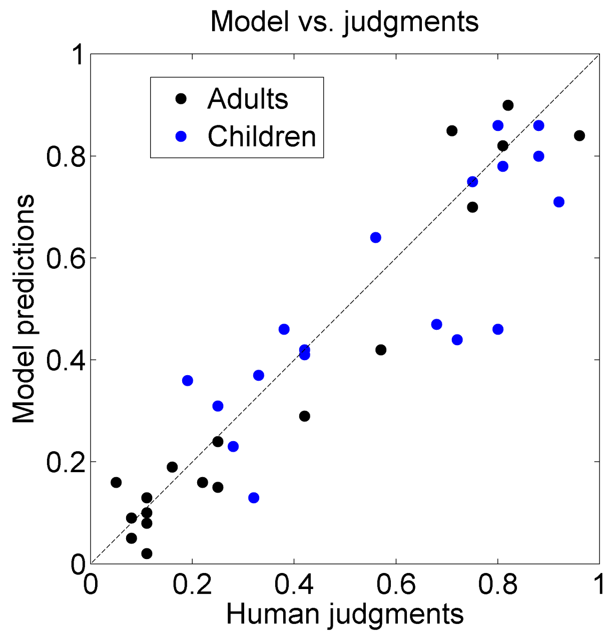

Figure 10. Model fits plotted against human judgments across all test objects (D, E, F, G) and all experiments. The overall correlation between the model's predictions and human judgments was r = 0.93. The dashed line represents a perfect correspondence between model predictions and human judgments.

Exploration, exploitation, and simulated annealing

Our empirical results and the model support the idea that children are quicker to learn the conjunctive relationship because they have more diffuse expectations than adults. However, there is another possible explanation for the developmental differences we have observed. It is possible that adults and children do not have different priors, but instead they react differently to new evidence. Children and adults update their beliefs in ways that reflect their distinct goals and circumstances.

Most of the time, adults do not need to dramatically change their beliefs or abandon their hypotheses for dramatically different ones. Indeed, doing so would be a liability: adults are expected to make accurate predictions and good decisions, not bold inductive leaps. Adults are also unlikely to have caregivers to correct their errors and save them from poor choices.

Children, on the other hand, face different challenges: they start from a position of relative ignorance and must revise their beliefs in fundamental ways, often engaging in radical conceptual change as they construct new intuitive theories (e.g., Gopnik & Meltzoff, 1997; Carey 1985, 2009). At the same time, children pay a much lower price for making incorrect decisions. Adults must exploit the knowledge they already possess. Children must explore the world around them and update their beliefs quickly.

The solution to this "explore-exploit trade-off" (e.g., Sutton & Barto, 1998) can be formalized in many different ways. One approach, which has a natural correspondence in Bayesian models, is based on the notion of *simulated annealing*. A popular algorithm for performing Bayesian inference, called Markov chain Monte Carlo (MCMC), explores a hypothesis space by proposing local, typically small-scale changes to existing hypotheses, and tending to accept proposals that are plausible and consistent with the available data (see Appendix B for further details). MCMC results in correct inferences in the long-run and explains some idiosyncrasies of human learning (e.g., Lieder, Griffiths & Goodman, 2012). However, in practice MCMC algorithms can be unacceptably slow to find good hypotheses, especially when the space of possibilities is very large.

One solution is to modify the process by which new hypotheses are assessed, so that the hypothesis space is explored more quickly. Simulated annealing is a method for doing this: it smooths out the probability distribution over hypotheses, so that lower-probability hypotheses are accepted more often. In the long run, this approach leads to incorrect inferences, because its exploratory approach makes it too quick to abandon good hypotheses. But it can be valuable in the early stages of learning, when the primary goal is to find reasonable hypotheses at a coarse level of granularity. Over time, the smoothing of the probability distribution is decreased, and the inference process shifts to a more conservative and asymptotically correct approach.

While this proposal is qualitatively different from the idea that children have more diffuse priors than adults, in practice it leads to virtually identical predictions in tasks such as ours. In many cases a higher-variance prior resembles a stronger inclination to explore new hypotheses. The key difference is that simulated annealing corresponds to having a more diffuse likelihood as well as a more diffuse prior, meaning that the learner is less influenced by each observation. We cannot distinguish between these two proposals on the basis of our current results, but future work, focusing on the dynamics of belief revision, may provide a direct test of the annealing proposal.

Conclusion

We have found that children can quickly learn about the forms that causal relationships take, and apply that knowledge to make judgments about new objects and to craft interventions. When the evidence indicates a conjunctive relationship is present, children learn and generalize more readily than adults. These results have implications for understanding causal learning and cognitive development more generally. In terms of causal learning, they suggest that abstract constraints that guide future inferences may themselves be learned (see also Kuhl, 2004 and Dewar & Xu, 2010). We believe that trying to understand the origins of these constraints is fertile ground for future research.

For cognitive development, the idea that children are more flexible in their commitments about causal systems may provide an important insight about the differences between children and adults. The very fact that children know less to begin with may, paradoxically, make them better, or at least more open-minded, learners. The plasticity of early beliefs may help to explain the bold exploration and breathtaking innovation that characterizes children's learning. Finally, our results suggest that a hierarchical Bayesian approach may help explain both how we reason as adults and how we learn as children.

Authors' Note

This research was supported by the James S. McDonnell Foundation's Causality Collaborative Initiative, Air Force Office of Scientific Research, grants FA9550-07-1-0351 and FA9550-10-1-0232 to T. L. G., and National Science Foundation grant (BCS-1023875) to A. G. and T. L. G. A preliminary version of this work was presented at the 32nd Annual Meeting of the Cognitive Science Society and appears in the proceedings of the conference (Lucas, Gopnik & Griffiths, 2010).

References

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, Mass.: MIT Press.

Carey, S. (2009). The origin of concepts. Oxford University Press, USA.

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

Dewar, K. M., & Xu, F. (2010). Induction, Overhypothesis, and the Origin of Abstract Knowledge Evidence From 9-Month-Old Infants. *Psychological Science*, 21(12), 1871-1877.

Dias, M. G., & Harris, P. L. (1990). The influence of the imagination on reasoning. British Journal of Developmental Psychology, 8, 305–318.

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review*, 7(4), 707–712.

Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge: Harvard University Press.

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

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

Gopnik, A., & Schulz, L. (2007). Causal Learning: Psychology, Philosophy, and Computation. Oxford University Press, USA.

Gopnik, A., & Wellman, A. (2012) A. Gopnik & H.M. Wellman (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin,* 138, 6, 1085-1108. doi: 10.1037/a0028044 1085-1108

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*, 354-384.

Gweon, H., & Schulz, L. (2011). 16-month-olds rationally infer causes of failed actions. Science, 332(6037), 1524.

Kemp, C., Perfors, A., & Tenenbaum, J. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*(3), 307–321.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience*, *5*(11), 831–843.

Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology*, 43(1), 186-196.

Lieder, F., Griffiths, T., & Goodman, N. (2012). Burn-in, bias, and the rationality of anchoring. In P. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25 (pp. 2699–2707).

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological review*, *115*(4), 955–984.

Lucas, C., & Griffiths, T. (2010). Learning the Form of Causal Relationships Using Hierarchical Bayesian Models. *Cognitive Science*, *34*(1), 113–147.

Lucas, C., Gopnik, A., & Griffiths, T. L. (2010). Developmental differences in learning the forms of causal relationships. In *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 28–52).

Markovits, H., & Vachon, R. (1990). Conditional reasoning, representation, and level of abstraction. *Developmental Psychology*, 26 (6), 942–951

Pascalis, O., Scott, L. S., Kelly, D., Shannon, R., Nicholson, E., Coleman, M., & Nelson, C. A. (2005). Plasticity of face processing in infancy. Proceedings of the National Academy of Sciences of the United States of America, 102(14), 5297–5300.

Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared cause tummy aches? Naive theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology*, *43*(5), 1124–1139.

Causal Learning Across Domains. Schulz, Laura E.; Gopnik, A. (2004) *Developmental Psychology*, 40(2) ,162-176

Schulz, L. E., Goodman, N. D., Tenenbaum, J. B., & Jenkins, A. C. (2008). Going beyond the evidence: Abstract laws and preschoolers' responses to anomalous data. Cognition, 109(2), 211–223.

Seiver, E., Gopnik, A. & Goodman, N. D. (2012). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development*. doi: 10.1111/j.1467-8624.2012.01865

Sobel, D. M., & Kirkham, N. Z. (2007). Bayes nets and babies: infants' developing statistical reasoning abilities and their representation of causal knowledge. *Developmental Science*, 10(3), 298-306.

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.

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. Cambridge University Press.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331*(6022), 1279-1285. doi: 10.1126/science.1192788

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition Without Control When a Little Frontal Lobe Goes a Long Way. Current Directions in Psychological Science, 18(5), 259–263.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, *43*, 337-375.

Xu, F. & Kushnir, T. (Eds.) (2013). *Rational Constructivism in Cognitive Development* Elsevier Inc.: Academic Press.

Appendix A: Model details

Following Lucas and Griffiths (2010), we use a logistic sigmoid family of likelihood functions to represent a continuous space of overhypotheses, where the probability of the machine's activation given that *n* blickets are present is

$$P(\text{effect}|N_{\text{blickets}} = n) = \frac{1}{1 + \exp\{-g(n-b)\}} \quad . \quad (1)$$

The overhypotheses determine the probability of different values of the bias *b* and the gain *g*. The bias specifies how many blickets are necessary to activate the machine, and the gain reflects how noisy the relationship is. Lucas and Griffiths found that exponential priors predicting a high mean gain (3.34) and a low mean bias (0.23) – or a reliable disjunctive relationship – lead to model predictions that closely match adults' judgments. If children are happier believing that a relationship could be conjunctive or noisy, the priors that best capture their inferences should lead to *a priori* gains and biases closer to 1. This space of likelihood functions is intended to cover a range of relationships that are appropriate to the cover story and participants' prior knowledge, and we do not claim it includes all relationships that people could conceivably learn.

Model fits.

We treated is-a-blicket judgments as assertions that objects were definitely blickets, and not-a-blicket judgments as assertions that objects were definitely not blickets. In order to test the proposal that children are more flexible learners than adults, we represented the priors over the gain and bias using gamma distributions, which generalize the exponential priors used in Lucas and Griffiths (2010). Gamma distributions can be parameterized by their means (μ) and variances (σ^2), which in our case reflect what relationships the learner expects and the strength of the learner's commitment to those expectations, respectively. After initial simulations that revealed broadly appropriate priors, we explored gain distributions with means from 1.0 to 7.0 and variances from .2 to 1.5, and bias distributions with means from 1.0 to 5.0 and variances from .05 to 2.0. For children, the best-fitting priors were μ_{gain}=1.2, σ^2_{gain}=1.3, μ_{bias}=1.2, and σ^2_{bias}=1.9. For adults, the best-fitting priors were μ_{gain}=3, σ^2_{gain}=5, μ_{bias}=.2, and σ^2_{bias}=.05. We also considered different priors over the probability that a novel object is a blicket, using beta distributions parameterized by the mean probability that a novel object is a blicket, and a virtual sample size. Simulations indicated that the mean parameter did not strongly influence fits for either group, with an optimum at approximately .5, and the best-fitting sample-size parameters were 9.0 for children and 2.0 for adults. Inference was performed using 500000 Metropolis-Hastings samples, which led to standard deviations on error estimates of less than .01.

Alternative models

To provide some context for the performance of our model, we also assessed the ability of alternative models to fit our data. One natural alternative is the best possible non-hierarchical model, that is, the best-fitting model that does not predict an influence of training-phase objects on test-phase judgments. Such a model has minimal error subject to the constraint that predictions for D, E, and F cannot vary between conditions. This is achieved by using as predictions the observed mean judgments for a particular age group and object, and using them across all conditions.

For our adult participants, such an approach yields similar performance to our own model, as measured by sum squared error (0.07 versus 0.08) and correlation (0.98 in both cases). This is not especially surprising, as adults' judgments varied relatively little between conditions. In contrast, this baseline model performs substantially worse than our own in predicting the judgments of children (sum squared error of 0.44 versus 0.17, correlation 0.72 versus 0.91). These fits provide upper bounds on the performance of any non-hierarchical model.

Another alternative is a model like ours, but which attempts to fit both groups using a single set of priors. We searched for good global priors by exploring 800 points in a volume in the parameter space that encompassed the best fits for the children and the adults, which varied the gain and bias terms as well as sensitivity to base rates, but not yes-bias. This model gave a sum squared error of 0.99, compared with 0.25 for the distinct priors we used.

Finally, we considered an alternative model that supposes that children might be making judgments like adults, but with added variability. Reflecting this assumption, we examined fits using the best adult priors after adding noise to the distribution of responses (bringing them closer to 0.5). The noise level that minimized error yielded a minimum sum squared error of 0.72 when compared to judgments from the children, still substantially higher than the 0.17 error that our model produced.

A systematic assessment of the roles of the free parameters is difficult, given that we cannot be certain that we have found optimal parameter values, a problem that would be compounded if we were to apply standard methods for compensating for possible over-fitting, such as cross-validation. Nonetheless, based on the dramatically better fits under our hierarchical model with distinct priors, as well as corroborating results from our inferential statistics, there seems to be ample evidence that the model compares favorably to the alternatives.

Appendix B: Simulated annealing

Simulated annealing is a technique for improving efficiency in solving problems of optimization, search, and inference. It takes its name from annealing in metallurgy and glassworking, where a material is heated and then slowly cooled, allowing it to reach lower-energy states that tend to have desirable properties. As in physical annealing, simulated annealing in a statistical setting depends on the concept of temperature, which determines how aggressively different hypotheses are explored. When the temperature is high, hypotheses are accepted regardless of whether or not they are supported by the data or likely a priori. This makes it possible to explore the space of possible hypotheses quickly, even when parts of that space are separated from the learner's initial beliefs by very unlikely hypotheses. While high-temperature search is useful for exploring different possibilities, it quickly abandons good hypotheses, so it is an unwise approach if one begins with hypotheses that are close to the most likely ones given the data. At the opposite extreme, very low temperatures lead to rejecting all hypotheses that are worse than the current hypothesis. This approach is useful for fine-tuning a single, firmly-held belief, but it cannot improve much on a poor starting hypothesis. By starting with a high temperature that decreases until hypotheses are accepted at a rate that reflects their true probabilities, simulated annealing has the advantages of high temperature search but still converges to good solutions in the long run. Such an approach is a natural fit for a learner like a child, who starts from a position of relative ignorance, is insulated from the consequences of poor decisions, and must ultimately solve a challenging problem in a very large space of hypotheses. It thus seems plausible that children approach inference in a manner that strongly resembles simulated annealing, and are quick to abandon old beliefs, good and bad, but become more attached to their hypotheses as they age.

Formally, the use of simulated annealing in a model like ours changes the probability that a new hypothesis h' will be accepted over a starting hypothesis h

$$P(accept) \propto \left(\frac{f(h')}{f(h)}\right)^{k}$$

where the k term increases over time (being inversely proportional to the temperate) until it is equal to 1.0, at which point the inference is identical to that found in a standard Metropolis-Hastings sampler. The function f is the un-normalized probability density of hypotheses given the data. This density is equal to the prior probability of the hypothesis times the likelihood of the data given the hypothesis, and one consequence of using a high temperature is that the prior term is effectively smoothed out: if the data are equally consistent with two hypotheses, having a high temperature is equivalent to having a less-concentrated prior distribution, consistent with the differences we observed between children and adults in our experiments.