PAPER

Little Bayesians or little Einsteins? Probability and explanatory virtue in children's inferences

Angie M. Johnston,[1,*] Samuel G.B. Johnson,[1,*] Marissa L. Koven[2] and Frank C. Keil[1]

1. Department of Psychology, Yale University, USA
2. Department of Psychology, Emory University, USA

Abstract

Like scientists, children seek ways to explain causal systems in the world. But are children scientists in the strict Bayesian tradition of maximizing posterior probability? Or do they attend to other explanatory considerations, as laypeople and scientists – such as Einstein – do? Four experiments support the latter possibility. In particular, we demonstrate in four experiments that 4- to 8-year-old children, like adults, have a robust latent scope bias *that leads to inferences that do not maximize posterior probability. When faced with two explanations equally consistent with observed data, where one explanation makes an unverified prediction, children consistently preferred the explanation that does* not *make this prediction (Experiment 1), even if the prior probabilities are identical (Experiment 3). Additional evidence suggests that this latent scope bias may result from the same explanatory strategies used by adults (Experiments 1 and 2), and can be attenuated by strong prior odds (Experiment 4). We argue that children, like adults, rely on 'explanatory virtues' in inference – a strategy that often leads to normative responses, but can also lead to systematic error.*

Research highlights

- In four studies, we test whether children, like adults, use 'explanatory virtues' in evidence-based inferences.
- We presented 4- to 8-year-old children with two explanations, equally consistent with the observed data, but where one explanation made an unverified prediction.
- Children consistently failed to maximize posterior probability, preferring explanations that did not make unverified predictions.
- This bias was overridden in the face of strong prior odds, suggesting that children, like adults, consider both probability and explanatory virtue when evaluating explanations.

Beauty is truth, truth beauty, – that is all
Ye know on earth, and all ye need to know.
(John Keats, 'Ode on a Grecian Urn', 1819)

Introduction

Likeliness and loveliness in explanation

Children are often characterized as budding scientists. In the first years of life, young children perform impressive inductive feats, managing to decipher the vocabulary and grammar of one or more natural languages, to carve the world up into useful categories, and to map the causal structure of the physical and social worlds. These

Address for correspondence: Angie Johnston or Samuel Johnson, 2 Hillhouse Ave., New Haven, CT 06520, USA; e-mails: angie.johnston@yale.edu or samuel.johnson@yale.edu
*The first two authors contributed equally to this work.

accomplishments are all the more remarkable because, unlike mature scientists, children must induce this knowledge without the benefit of formal education or scientific training.

If children truly approach the world like little scientists, gathering evidence and inferring regularities, then perhaps their inferential practices are also similar to those of actual scientists. In order for scientists to make sense of the world, they must perform *abduction* – inferring the best explanation for a set of observations (Lipton, 2004). However, within philosophy of science, there is disagreement about what criteria scientists use for evaluating explanations. According to Bayesian confirmation theory (e.g. Jeffrey, 1965), science is concerned with inferring the *likeliest* explanation – the hypothesis that has maximum posterior probability after observing the evidence. On the reasonable assumption that seeking truth requires us to seek the most probable explanation, scientists certainly seem to aspire to this goal.

However, scientists may not directly consider which explanations are most likely in the sense of maximizing posterior probability, but may instead search for the *loveliest* explanation, in the hope that their intuitive sense of explanatory virtue can be a guide to truth (Lipton, 2004; McGrew, 2003). There is much anecdotal support for the importance of explanatory elegance to the work of scientists. For example, Hermann Bondi describes his experience meeting Albert Einstein (quoted in Zee, 1999):

> What I remember most clearly was that when I put down a suggestion that seemed to me cogent and reasonable, Einstein did not in the least contest this, but he only said, 'Oh, how ugly.' As soon as an equation seemed to him to be ugly, he really rather lost interest in it and could not understand why somebody else was willing to spend much time on it. He was quite convinced that beauty was a guiding principle in the search for important results in theoretical physics.

This remark is representative of a common sentiment among practicing scientists. Einstein is joined by legions of great scientists and mathematicians in echoing Keats' refrain: that 'beauty is truth, truth beauty'.

Laypeople appear to share Einstein's intuition that *likeliness* in explanation is not enough. Adults rely on a set of overarching 'explanatory virtues' (e.g. simplicity, scope, and generality) to inform their explanatory preferences (Lombrozo, 2012). Although these virtues often lead to inferences that maximize posterior probability (e.g. simpler explanations are often likelier than complex explanations), there are particular cases in which they do not (McGrew, 2003; van Fraassen, 1989). For example, adults are influenced by the *relative* quality of an explanation. Rather than assigning explanations weight based solely on the magnitude of their posterior probabilities, adults place extra weight on an explanation when it outperforms its competitors (Douven & Shupbach, 2015; Pacer, Williams, Chen, Lombrozo & Griffiths, 2013) or maximizes the likelihood of the data (Pacer *et al.*, 2013). In fact, adults sometimes place so much additional weight on these explanatory virtues that they prefer an explanation that has a *lower* probability of being true based on the data. For instance, adults prefer simpler explanations (invoking fewer causes) over more complex explanations (invoking more causes), even when a complex explanation has a higher probability of being true (Lombrozo, 2007). Thus, like Einstein, adults seem to be influenced by the *loveliness* of an explanation, even to the detriment of its *likeliness*.

Developmental origins of virtue-based explanatory inference

Although adults integrate explanatory virtues into their inferences over and above probability, it is unclear whether children do the same. If explanatory virtues are acquired over a long period of development as inferential 'short-cuts', then young children may be less likely to rely on virtues, and more likely than adults to focus on maximizing posterior probability. In contrast, if explanatory virtues are present from early in development, then children, like adults, should readily attend to these explanatory virtues, leading to inferences that do not maximize posterior probability.

Children are early consumers of explanations (e.g. Legare & Lombrozo, 2014; Frazier, Gelman & Wellman, 2009; Walker, Lombrozo, Legare & Gopnik, 2014; Wellman, 2011), and like adults (Cimpian & Steinberg, 2014; Lombozo, 2007; Rips, 2002), use a variety of criteria for evaluating them. For example, young children prefer explanations that avoid circularity (Corriveau & Kurkul, 2014; Mercier, Bernard & Clément, 2014), provide detailed causal information (Frazier, Gelman & Wellman, 2016), invoke fewer causes (Bonawitz & Lombrozo, 2012), and implicate inherent properties (Cimpian & Steinberg, 2014; Hussak & Cimpian, 2015).

But it is less clear whether children use explanatory virtues over and above probability, as adults do. Young children are adept probabilistic reasoners (e.g. Gopnik, Glymour, Sobel, Schulz, Kushnir *et al.*, 2004; Gweon, Tenenbaum & Schulz, 2010; Schulz & Bonawitz, 2007; Schulz, Bonawitz & Griffiths, 2007), and only one developmental study has pitted posterior probability against explanatory virtue (Bonawitz & Lombrozo, 2012). In this study, 4- to 6-year-old children were asked to evaluate explanations that differed in (a) their prior probabilities and (b) the number of causes they invoked. Children were introduced to a machine that had a light

and a fan. Red coins caused the light to turn on, green coins caused the fan to turn on, and blue coins caused both the fan and the light to turn on. The experimenter 'accidentally' tipped a bag of coins over, causing both the light and fan to activate. Thus, either a blue coin (causing the fan and light to turn on), or both a green coin (causing the fan to turn on) and a red coin (causing the light to turn on), must have fallen into the machine. Children believed that the simple (blue coin) explanation was more likely than the complex explanation (green and red coins), even if there were many red and green coins but only one blue coin. In fact, it took odds favoring the complex explanation by four or six times to override this preference. Thus, like adults, children favor the explanatory virtue of simplicity, failing to infer the explanation with maximal posterior probability.

The latent scope bias

To further investigate whether children are influenced both by explanatory virtues and probabilistic information, we capitalized on an explanatory bias shown by adults: The *latent scope bias*, wherein explanations that make fewer unverified or latent predictions are preferred (Johnson, Rajeev-Kumar & Keil, 2014c, 2016; Khemlani, Sussman & Oppenheimer, 2011). For example, imagine that your car smelled like antifreeze, and this could be due to one of two equally common problems – a problem with the cooling system or a problem with the exhaust. Suppose that a cooling problem would activate the engine light, but an exhaust problem would not. Clearly, the thing to do is to check the light. But alas, the light is useless because the bulb has burned out! In this situation, the light is in the *latent scope* of the cooling system explanation – that is, the light would count as evidence in favor of a cooling problem if it were observed, but the prediction is unverified. Based on the posterior probabilities of the two explanations, both are equally likely. Yet, adults tend to prefer explanations with fewer latent effects and thus would say that the exhaust explanation – which does not predict any additional effects – is more satisfying and more probable.

This inference in adults appears to result from a combination of two explanatory processes (Johnson *et al.*, 2014c, 2016): *inferred evidence* and *manifest scope*.

Inferred evidence

First, when confronted with an explanation that makes an unverified prediction, adults make an inference about the latent evidence to resolve this ignorance, effectively guessing whether the evidence would be observed if they were able to see it. In doing so, adults rely on the base rates of the unknown effect, even if the prior probabilities of the explanations (the only relevant information for determining the posterior probabilities; see Johnson *et al.*, 2016) are explicit in the problem. For example, in the case of the antifreeze smell, an adult would first infer that the engine light is probably off, because the engine light is off most of the time.

Manifest scope

After making an inference about the latent evidence, adults evaluate competing explanations based on the evidence they have inferred. In the antifreeze example above, if adults infer that the engine light is off, then they would evaluate the competing explanations based on the evidence that (a) it smells like antifreeze and (b) the engine light is off. In this case of the latent scope bias, adults would apply a *manifest scope* preference, preferring explanations that account for as many confirmed and as few disconfirmed observations as possible (Johnson, Johnston, Toig & Keil, 2014b; Read & Marcus-Newhall, 1993).

Given that a cooling problem would have activated the light (but an exhaust problem would not), the exhaust problem is a better explanation, since it directly matches the inferred evidence. That is, the data are more likely under the exhaust explanation than under the cooling explanation, yielding a higher posterior probability. Thus, in contrast to the inferred evidence step, the manifest scope step directly relates to the posterior likelihood of the explanation. Its use only leads to non-maximizing behavior in the case of the latent scope bias because the inferred evidence does not accurately reflect the base rates of the explanations.

Developmental precedents

Although there are no direct tests of a latent scope bias in children, there is some indirect evidence that young children might use each of its two component processes (i.e. inferred evidence and manifest scope).

First, young children might reason about inferred evidence in cases where they are reluctant to accept epistemic ignorance. For instance, when trying to determine which of two locations an object might be hidden in, 4- to 6-year-olds are willing to say that either hiding location is possible when uncertainty resides in the physical world (because the object has not yet been hidden; Robinson, Rowley, Beck, Carroll & Apperly, 2006). However, when uncertainty resides in their mind (because the object has already been hidden), children

frequently guess a particular hiding location – even though they have no way of knowing where the object is (Robinson *et al.*, 2006). This suggests that children are highly motivated to resolve epistemic ignorance, even guessing arbitrarily to do so. This motivation could potentially lead children to make inferences about unknown explanatory evidence in the same way as adults.

Children may also have a manifest scope preference. By age 4, they are more likely to privilege causes that account for a greater number of prior observations when making causal predictions (Walker, Williams, Lombrozo & Gopnik, 2012), and by age 7, children prefer explanations that account for a wider range of data points (Samarapungavan, 1992). In addition, 4-year-olds, like adults, have a robust preference for simpler explanations that invoke fewer causes (Bonawitz & Lombrozo, 2012; Lombrozo, 2007). Because children believe that every event has a cause (Bullock, Gelman & Baillargeon, 1982; Schulz & Sommerville, 2006), explanations that only explain a subset of the available data would need to be conjoined with additional explanations to explain the other data. Thus, young children's preference for simpler explanations is indirect evidence for a preference for wide manifest scope.

Thus, it seems that even young children have relevant cognitive elements that might lead them to share adults' bias against explanations that make unverified predictions. Like adults, they (a) are motivated to resolve epistemic ignorance about unobservable information (Robinson *et al.*, 2006), (b) privilege causes that account for a greater number of observations when making causal predictions (Walker *et al.*, 2012), and (c) prefer explanations that invoke fewer causes (Bonawitz & Lombrozo, 2012). The current experiments test children's explanatory inferences about latent observations more directly.

The current studies

Four experiments investigate whether children, in the face of latent observations, maximize posterior probability, or instead incorporate explanatory virtues as adults do. If children maximize the posterior, they should be unconcerned about unverified predictions because they are not relevant. Instead, they should, perhaps tacitly, (1) calculate the prior probabilities of both explanations and their ratio (i.e. the prior odds), (2) calculate the likelihoods of both explanations and their ratio (i.e. how probable the data would be under each hypothesis; when all that varies across explanations is latent scope, this ratio is 1 for deterministic causal systems, because the known evidence is predicted by

both hypotheses), and then (3) multiply these two ratios (as dictated by Bayes rule; Pearl, 1988). In contrast, if children go beyond a narrow consideration of posterior probability and consider a set of overarching explanatory virtues, they might show evidence of the same latent scope bias as adults.

Across four experiments, we investigate whether children show the same latent scope bias as adults. In Experiments 1 and 3 we test whether 5–8-year-olds (Experiment 1) and 4–5-year-olds (Experiment 3) show an adult-like preference for narrow latent scope, preferring explanations that do not make unverified predictions. We build on these results in Experiment 4 by varying the base rates to test whether this preference is sensitive to probability as simplicity is (Bonawitz & Lombrozo, 2012). Although our primary concern is whether children show a latent scope bias overall, we also begin to examine whether the underlying mechanisms of this bias may be the same in childhood as they are in adulthood. To this end, we examine each of the two component processes of the adult latent scope bias – manifest scope (Experiment 1) and inferred evidence (Experiment 2) – separately. We anticipated across these experiments that young children would use adult-like explanatory reasoning, leading them to go beyond a narrow consideration of posterior probabilities in order to consider explanatory virtue as well.

Experiment 1

In our first experiment, we asked whether children, like adults, prefer explanations that fully account for the evidence (wide *manifest scope*) and do not make unverified predictions (narrow *latent scope*). If children share the same explanatory virtues as adults, then they should show both preferences. In contrast, if they maximize the posterior probability based on the observations, they should not show a latent scope preference.

Methods

Participants

Sixty-four 5- to 8-year-old children, divided evenly between the younger (5–6) and older (7–8) halves of the age group, participated. Ten children (seven 5- to 6-year-olds and three 7- to 8-year-olds) were excluded because they gave inconsistent responses (see below). The final sample included 25 5- to 6-year-olds ($M = 5$ years, 11 months; range = 5 years, 0 months–6 years, 11 months; 10 females) and 29 7- to 8-year-olds ($M = 7$ years, 11 months; range = 7 years, 0 months–8 years, 10 months; 19 females).

Materials and procedure

Each child completed four items. Because previous studies have found domain differences in explanatory preferences (e.g. Johnston, Sheskin, Johnson & Keil, 2016), half of the children evaluated biology explanations about animals and the other half evaluated physics explanations about machines (see Supplementary Materials for sample stimuli).

For each item, the experimenter read a story to the child, which was accompanied by animated illustrations in Microsoft Powerpoint. For example, for one item, a pig appeared on a screen in full view, so that the child could see that the pig had typical features. Then, the experimenter pressed a button, and the pig moved across the screen, passing behind a tree and emerging partially, so that its head and body could be seen, but its tail was occluded by the tree. The experimenter told the child:

> This pig went behind a tree. When it went behind the tree, it accidentally ate a special acorn, but we don't know which one.

The experimenter then told the child about three types of acorns that the pig could have eaten (in a pseudorandomized order, explained below):

> *1-effect*: Blue acorns make pigs get stripes on their ears.
> *2-effect*: Purple acorns make pigs get stripes on their ears and make them grow whiskers.
> *3-effect*: Green acorns make pigs get stripes on their ears, make them grow whiskers, and make their tails uncurl.

As each acorn was described, a pictorial diagram was placed in front of the child to summarize its effects. After describing the possible explanations, the experimenter pressed a button so that the changes could be observed in real time. As each change occurred, the experimenter narrated what was happening:

> When the pig went behind the tree, it got stripes on its ears and grew whiskers. Looks like we can't see anything else, but at least we know it got stripes on its ears and it grew whiskers.

The pig's tail was occluded by the tree that it had passed behind, making any predictions involving the tail unverifiable, and thus part of the latent scope (a method we adopted from Sussman, Khemlani & Oppenheimer, 2014). Importantly, the experimenter did not explicitly mention the tail, to avoid the possibility that drawing the child's attention to the absence of evidence might pragmatically imply evidence of absence.

Given these observations (stripes and whiskers, but no information about the tail), the 1-effect explanation has narrow manifest scope because it accounts for only one of two actual observations, and the 2-effect explanation has wide manifest scope because it accounts for both observations. We would therefore expect a preference for the 2-effect over the 1-effect explanation if children prefer wide manifest scope, as adults do (Johnson *et al.*, 2014b; Read & Marcus-Newhall, 1993). With respect to latent scope, the 3-effect explanation has wide latent scope because it accounts for both actual observations but also makes an unverified prediction, and the 2-effect explanation has narrow latent scope because it does not make any unverified predictions. If children have a narrow latent scope preference, like adults (Khemlani *et al.*, 2011), then they should prefer the 2-effect explanation over the 3-effect explanation.

For each item, two types of questions were asked. The first questions tested *hypothesis pruning* – determining which potential explanations are worth entertaining as possibilities (Johnson & Keil, 2014; Peirce, 1997/1903). These questions tested whether the child used narrow manifest scope or wide latent scope to narrow the possibility space. The second type of question tested *hypothesis evaluation* – determining which out of the candidates deemed possible is considered best (Khemlani *et al.*, 2011; Lombrozo, 2007). Together, these two questions help to address how explanatory virtues are applied at a process level: Do children prune out all potential explanations except those that account for all and only the existing data, or do they consider a wider variety of explanations and then select the best explanation among the candidates?

The first three questions tested hypothesis pruning. For each explanation, the experimenter pointed to its card and asked, 'If the pig only ate one kind of acorn, could it have eaten the [*color*] acorn to make it get stripes on its ears?' These yes/no questions were asked for the three explanations in the same order that they were introduced during the story. The fourth question for each item tested hypothesis evaluation. The experimenter asked, 'Which kind of acorn do you think it ate to make it get stripes on its ears?' Importantly, the experimenter only asked about *one* of the two observed effects (e.g. referring to the 'stripes', rather than both the 'stripes and whiskers') in all questions. This required children to notice the relevance of the second observed effect on their own. Ten children were excluded from analysis because their explanation choice (in the second question) was not among the explanations they deemed possible (in the first question) for at least one of the four questions.

The order of the four items (e.g. pig, lizard, dog, lion) and the order of the explanations within each item were counterbalanced, as was whether the latent effect was mentioned first or last in the 3-effect explanation.

Results

Preliminary analyses revealed no main effects or interactions involving age, domain, or the counterbalancing factors, so we collapsed across these factors.

Pruning questions

First, we investigated whether children were any more or less likely to prune explanations from the field of possibilities based on whether they were 1-effect (narrow manifest scope), 2-effect (wide manifest scope but narrow latent scope), or 3-effect (wide latent scope). As shown in Figure 1, the 2-effect explanations were seldom pruned, as children included these explanations in the hypothesis space on nearly all trials ($M = 3.48$, $SD = .91$ out of 4). In contrast, paired-samples t-tests revealed that, relative to the 2-effect explanations, children were more likely to eliminate both the 1-effect ($M = 2.50$, $SD = 1.55$), $t(53) = 4.70$, $p < .001$, $d = 0.77$, and 3-effect explanations ($M = 2.39$, $SD = 1.62$), $t(53) = 4.90$, $p <$
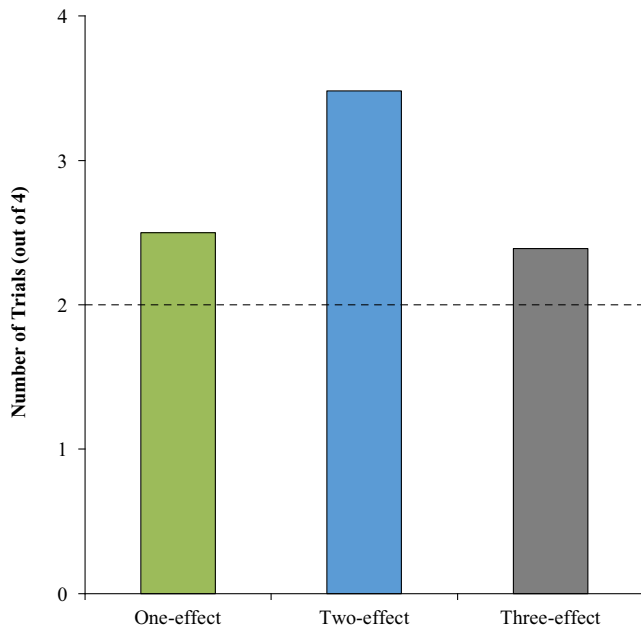
.001, $d = 0.83$. However, regardless of explanation type, children were more likely than chance to include explanations in the possibility space ($p = .022$ for 1-effect; $p < .001$ for 2-effect; $p = .083$ for 3-effect), although this tendency to include the 3-effect explanation with wide latent scope was marginal.

Evaluation questions

The primary question of interest was whether children would prefer the 1-effect (narrow manifest scope), 2-effect (wide manifest scope but narrow latent scope), or 3-effect (wide latent scope) explanation in the evaluation questions. As children were presented with a forced-choice between the three levels of explanation, they could select each explanation as many as 4 times, or as few as 0. As shown in Figure 2, children preferred the 2-effect explanation ($M = 2.59$, $SD = 1.31$) more than either the 1-effect ($M = 0.85$, $SD = 1.00$), $t(53) = 6.06$, $p < .001$, $d = 1.49$, or 3-effect explanations ($M = 0.56$, $SD = .98$), $t(53) = 7.16$, $p < .001$, $d = 1.76$. There was no difference in their preference for the 1-effect and 3-effect explanations, $t(53) = 1.46$, $p = .149$, $d = 0.29$. Moreover, children selected the 2-effect explanation more than chance, $t(53) = 7.06$, $p < .001$, $d = 1.94$, and selected both the 1-effect and 3-effect explanations less than chance ($ts > 3.54$, $ps < .002$). Thus,
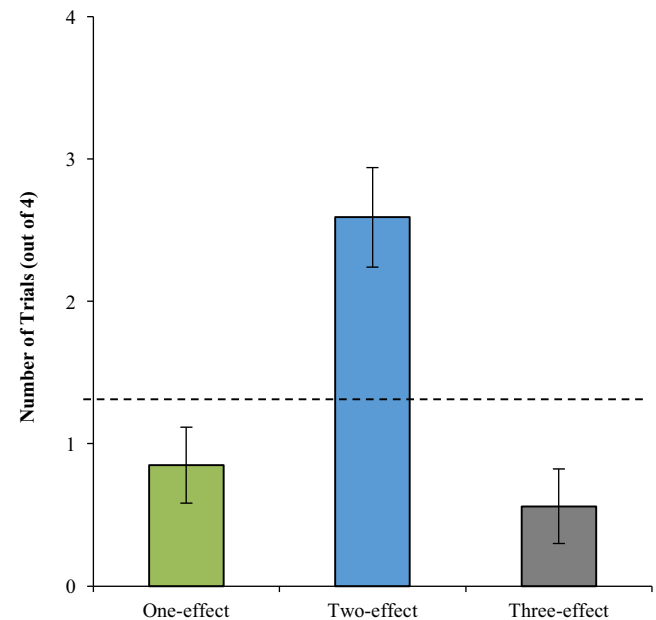

Figure 1   *Pruning results from Experiment 1. Bars depict the number of trials (out of 4) in which an explanation was included in the hypothesis space (as a yes/no choice for each explanation). Dashed line indicates chance responding (50%), and error bars represent 95% CIs.*

Figure 2   *Evaluation results from Experiment 1. Bars depict the number of trials (out of 4) in which an explanation was selected (as a forced choice among the three explanations). Dashed line indicates chance responding (33%), and error bars represent 95% CIs.*

like adults, children prefer explanations with wide manifest scope – explaining more of what is observed – and narrow latent scope – predicting little that is not known to be true.

Discussion

These results suggest that children as young as age 5 use explanatory scope in a similar way to adults. Although most children thought at some level that all three explanations were possible, they were nonetheless far more likely to prune the 1- and 3-effect explanations than the 2-effect explanation from the hypothesis space, and had a strong preference for the 2-effect explanation in a forced choice. The preference for the 2- over the 1-effect explanation is consistent with children's preference for simple explanations (Bonawitz & Lombrozo, 2012) and demonstrates that children have a preference for wide manifest scope – explanations that account for as many observations as possible. The preference for the 2- over the 3-effect explanation demonstrates that children have a preference for narrow latent scope – explanations that do not make unverified predictions (Khemlani *et al.*, 2011). Although the *wide manifest scope* preference is consistent with maximizing posterior probability, the *narrow latent scope* preference is not: Children have no direct evidence one way or another regarding whether the latent effect occurred, so they have equal posteriors (Johnson *et al.*, 2016). Thus, children seek explanatory virtues even when probability information does not discern between the two options.

Experiment 2

In adults, the latent scope bias is a product of two processes: manifest scope and inferred evidence. Whereas Experiment 1 provided direct evidence for a manifest scope preference, Experiment 2 tested whether children use inferred evidence about unverifiable predictions when evaluating explanations.

As in Experiment 1, we presented children with events in which two effects were observed and one was unknown (because it was occluded from view). However, in Experiment 2, children were only presented with two explanations – each of which made a different prediction about the unknown effect. Since both explanations made unverified predictions, they both had *wide latent scope*. By keeping the latent scope of the explanations constant, we were able to directly examine the inferred evidence component of the latent scope bias.

For each item, we first asked children to make an inference about the unknown effect and then asked children to select the best explanation for one of the *observed* effects. Based on prior work (Robinson *et al.*, 2006), we anticipated that children would guess arbitrarily when making their initial inference. However, the question of interest was whether these arbitrary guesses would connect to children's explanatory inferences. If children incorporate inferred evidence into their explanatory evaluations, then they should prefer explanations that line up with the evidence they inferred – even when (1) their inference was arbitrary (because the evidence was occluded) and (2) they are asked to explain a different observation.

Method

Participants

Twenty 5- to 6-year-olds (M = 5 years, 11 months; range = 5 years, 2 months–6 years, 7 months; 9 females) and 20 7- to 8-year-olds (M = 7 years, 10 months; range = 7 years, 0 months–8 years, 10 months; 8 females) participated.

Materials and procedure

Children completed two items adapted from Experiment 1 – the pig and the lizard (see Supplementary Materials for sample stimuli). For example, for one item, a pig appeared on the screen with its tail occluded by a tree that had five green acorns and five purple acorns on its branches. After allowing the child to briefly look at the image, the experimenter pointed out that there were an equal number of green acorns and purple acorns on the tree (establishing that the base rates of two types of acorn were equal). Then, the experimenter told the child what each type of acorn would do if the pig ate it. Both acorns predicted that the pig would get stripes on its ears and grow whiskers, but the two explanations differed in their predictions about the latent effect – one explanation predicted a curly tail and one predicted a straight tail. As in Experiment 1, a physical diagram was placed in front of the child to summarize the effects associated with each acorn. The order of the items and explanations was randomized across participants.

After describing the two potential explanations, the experimenter told the child that the pig covered its eyes and ate one of the acorns, emphasizing that the pig could not see which acorn it ate. This was done to ensure that children did not try to reason about which acorn the pig would *want* to eat. Then, the child saw an animation of what happened to the pig when it ate the acorn. As in Experiment 1, two of the three predicted effects were observed and one effect was unknown (i.e. the tail), since it remained occluded from view.

For each item, two questions were asked. The first question tested *inferred evidence* – asking children to infer the state of the tail ('Do you think the pig has a straight tail or a curly tail right now?'). When this question was asked, the experimenter pointed to the picture of the pig on the computer screen and asked children to verbally provide their answer. The second question asked children to make their *evaluation* about which explanation was best. As in Experiment 1, the experimenter only asked about *one* of the two observed effects ('Which acorn do you think the pig ate to make it get stripes on its ears?') and did not mention the unknown effect. This required children to notice the relevance of the unknown effect, and thus apply their inferred evidence, on their own.

Results

Children's *inferences* about the unknown effects were split equally across the two options for both items: out of 40 children, 20 children said the lizard's tongue was red (as opposed to blue), and 21 children said the pig's tail was curly (as opposed to straight). However, even though these inferences were arbitrary, children used them to evaluate the explanations. When children were asked to select an explanation for the first *observed* effect (e.g. 'Which acorn did the pig eat to make it get stripes on its ears?'), they tended to select the explanation that predicted the third *latent* effect they had previously inferred (e.g. a curly tail; $M = 1.60$ out of 2 items, $SD = .63$), $t(39) = 6.00$, $p < .001$, $d = 0.95$. There was no difference between the younger children ($M = 1.50$, $SD = 0.61$) and older children ($M = 1.70$, $SD = 0.66$), $t(38) = 1.00$, $p = .32$, $d = 0.32$.

Discussion

The results of Experiment 2 demonstrate that children's explanatory preferences (for wide manifest scope and narrow latent scope, as demonstrated in Experiment 1) are related to the inferences they draw about unverified evidence. Although children arbitrarily guessed the state of the unknown features (e.g. whether the pig's tail was curly or straight) in Experiment 2, they consistently preferred explanations that lined up with their inferences about the unknown features. This provides some initial evidence that children's latent scope bias (demonstrated in Experiment 1) may be driven by the same two explanatory processes as adults' latent scope bias: inferred evidence and manifest scope (Johnson *et al.*, 2016).

Although Experiment 2 sheds additional light on the way in which children evaluate competing explanations

with *wide* latent scope, there are still alternative explanations for children's *narrow* latent scope preference in Experiment 1. First, perhaps children in Experiment 1 interpreted their failure to observe the latent evidence as evidence that the effect did *not* occur, either because they were making pragmatic inferences on the basis of the evidence that was presented (e.g. Bonawitz, Shafto, Gweon, Goodman, Spelke *et al.*, 2011) or because they believed that occluded events could not occur. Second, perhaps children believed that explanations with narrow latent scope have higher base rates than explanations with wide latent scope (e.g. if children have the assumption that causes with fewer effects occur more frequently), given that they were never explicitly told about the base rates in Experiment 1. We tested these possibilities in Experiments 3 and 4.

Experiment 3

In Experiment 3, we tested for a latent scope bias using a task with transparent prior probabilities and likelihoods – a toy with a fan and light, similar to that used by Bonawitz and Lombrozo (2012). Children learned that one color coin turned on the fan (the *1-effect* coin) and that the other color coin turned on both the fan and light (the *2-effect* coin). After several familiarization trials with these coins (in which various parts of the toy were occluded in order to break potential pragmatic inferences about occlusion – e.g. that the experimenter was deliberately hiding evidence), children were presented with one test trial in which the light was occluded so that they could not tell whether it was on or not. Then, one coin was randomly and covertly put into the machine and children were asked to infer which coin was placed inside. The coin was drawn from a bag containing five coins of each color, to ensure that the prior probabilities were equal. If children are simply focused on maximizing the posterior probability, they should guess at chance because the fan is not diagnostic (it is consistent with either explanation), and the key piece of information (the light) is unavailable. In contrast, if children show a latent scope bias like adults (Khemlani *et al.*, 2011), they should indicate that the 1-effect coin is more likely since it does not make the additional, unverified prediction that the light would be on.

Method

Participants

Thirty-one 4- and 5-year-old children ($M = 4$ years, 11 months; range = 4 years, 0 months–6 years, 0 months; 16

females) participated. An additional 14 children (11 4-year-olds and 3 5-year-olds) participated but were replaced because they failed the familiarization check questions (see below).

Materials

The materials included a machine toy (see Figure 3), constructed from white cardboard. On the top of the machine, facing the child, were a fan that could rotate and a light that could turn on. A slot at the front of the machine was used to drop coins in, which purportedly caused the fan or light to operate. In fact, the fan and light were covertly operated by the experimenter using switches wired to the back of the box, out of view of the child. No child voiced suspicion over the operation of the machine; in fact, a senior museum staff member at one of our testing sites was surprised to learn that the coins did not control the machine.

Procedure

The procedure involved three phases: The *introduction*, *familiarization*, and *test* phases.

In the *introduction* phase, the experimenter explained the function of the blue and red coins. One coin (the *1-effect* coin) made just the fan turn on, while the other coin (the *2-effect* coin) made both the fan and light turn on. The color of the coins was counterbalanced, such that the 1-effect coin was blue for some children and red for others. For each coin, the experimenter put the coin
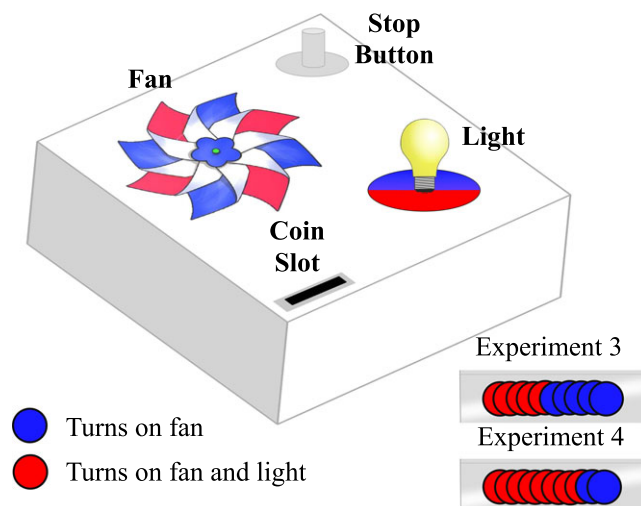

Figure 3 *Machine toy used in Experiments 3 and 4, including the coins that operated the machine and their base rates across experiments. The light was occluded on test trials so that children could not observe whether it was on. The toy was oriented so that the child faced the coin slot.*

in the slot so the child could witness what the coin caused the machine to do. The experimenter then said, 'See! The blue [*red*] coin makes the fan [*both the fan and the light*] go.' After introducing each coin, the experimenter gave a card to the child depicting the coin's color and its effects to reduce the task's memory load. The order in which the experimenter introduced the coins was randomized.

Next, in the *familiarization* phase, the child made six predictions – two in which both parts of the toy were visible and four in which one part was occluded – about what would happen if coins were put into the toy. If a child required more than one correction on the same familiarization trial (either visible or occluded), that child did not proceed to the test phase and was excluded from analysis. On the first set of familiarization trials (i.e. the two *visible* trials), the child was asked to predict what would happen when the red and blue coins were put into the slot. These trials were intended to make sure that the children understood how the machine worked. If the child answered incorrectly, the experimenter put the coin in to demonstrate the correct answer, and the trial was repeated. The order of the two visible trials (for the red and blue coins) was randomized.

On the second set of familiarization trials (i.e. the four *occluded* trials), either the fan or the light was covered up using an opaque cover, and the child was asked to predict what would occur when each coin was placed in the slot. These trials were framed as a guessing game, wherein parts of the machine were sometimes covered. This was done in order to break any pedagogical or pragmatic inferences children might be making about what the experimenter was communicating by covering the fan and light, and to ensure that children understood that unobserved effects could still occur. If the child answered incorrectly, the experimenter lifted the cover, and the trial was repeated. The order of the four invisible trials (for the red and blue coins, and with either the fan or light covered) was randomized.

Finally, in the *test* phase, the light was occluded. The test trial was continuous with the familiarization trials, so that from the child's perspective, covering the light on this trial was no different from covering parts of the machine on the previous familiarization trials. The experimenter showed the child a transparent plastic bag containing five red coins and five blue coins and said:

> We're going to use this bag of coins! See, there are five red coins and five blue coins in this bag. I'm going to close my eyes and pull one out. Then, I'll put it in the box, and I want you to guess which color went in.

Then, the experimenter and child both closed their eyes, and the experimenter selected a coin at random

from the bag, so that the child could not see which coin was selected. The experimenter then placed the coin in the slot and the appropriate effects occurred (i.e. the fan always turned on, and the occluded light did or did not turn on, depending on the coin color). Then, the experimenter asked, 'Which color do you think went in?' If children are averse to latent scope, they should choose the 1-effect option (the light should be off), but if they prefer latent scope explanations, they should choose the 2-effect option (the light should be on). Alternatively, if children are indifferent to latent scope and simply focus on maximizing the posterior probability, they should choose the coins equally often.

Results and discussion

As shown in Figure 4, children preferred the explanation with narrow latent scope – the coin that caused only the fan to turn on. Specifically, on the test trials, 24 out of 31 children (77%) chose the narrow latent scope coin ($p = .003$, sign test). This preference was equally strong among 4- and 5-year-olds ($p = 1.00$, Fisher's exact test). These results demonstrate that children as young as 4 have a robust latent scope bias, placing weight on some of the same explanatory virtues as adults (e.g. Johnson *et al.*, 2016; Khemlani *et al.*, 2011).
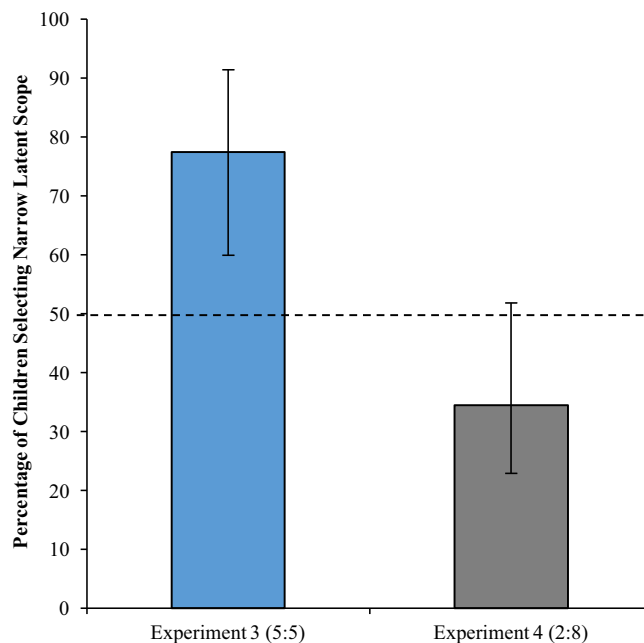

Figure 4   *Results of Experiments 3 and 4. Bars depict the percentage of children in each experiment who selected the narrow latent scope explanation. Dashed line indicates chance responding (50%), and error bars represent 95% CIs (using the method of Agresti & Coull, 1998, for binomial proportions).*

These results address several concerns about Experiment 1. The framing of the experiment – where parts of the machine were frequently covered up – would block any tendency children might have to interpret the cover as a communicative act on the part of the experimenter. Further, our check questions on the familiarization trials ensured that children understood that covered up events could still occur. Finally, the bag of coins set the prior probabilities of each explanation at precisely 50%, ruling out the possibility that children interpreted the base rates of the causes as different.

Experiment 4

Children have surprisingly sophisticated probabilistic reasoning skills, starting from infancy (Gweon *et al.*, 2010). In particular, children use the base rates of explanations to calibrate their preference for simpler over complex explanations (Bonawitz & Lombrozo, 2012). When the base rates of simple and complex explanations are made equal by varying the number of colored coins, children (like adults; Lombrozo, 2007) prefer the simple explanation. But when the complex explanation is much more probable than the simple explanation (e.g. a 1:6 ratio), children are able to override their simplicity preference. This ability is important because it allows the reasoner to attend to multiple cues. Would children similarly be able to override their latent scope bias when the base rates favor the wide latent scope explanation?

To test whether children are influenced both by explanatory scope and base rates when making explanatory inferences, we manipulated the prior odds using the method of Bonawitz and Lombrozo (2012). Instead of drawing a coin at random out of a bag with five 2-effect and five 1-effect coins as in Experiment 3, the bag contained eight 2-effect and two 1-effect coins. That is, the wide latent scope explanation had a prior probability that was 4 times higher than the narrow latent scope explanation. If children can consider probabilistic information alongside explanatory virtue, then this base rate manipulation should significantly weaken their preference for the 1-effect coins by making the 2-effect coins more probable. However, if overwhelming prior odds are still insufficient to override the latent scope bias, then they should continue to choose the 1-effect coins with narrow latent scope.

Method

Participants

Thirty-two 4- and 5-year-old children ($M = 4$ years, 11 months; range = 3 years, 11 months–5 years, 10 months;

15 females) participated. An additional 6 children (all 4-year-olds) participated but were replaced because they failed the same familiarization trial at least two times (the same criterion used in Experiment 3).

Materials and procedure

The materials and procedure were identical to those for Experiment 3, except for the test trial. On the test trial in Experiment 4, the experimenter used a bag of coins with eight 2-effect coins (i.e. wide latent scope) and two 1-effect coins (i.e. narrow latent scope), in contrast to Experiment 3 where five of each type of coin were used.

Results and discussion

As shown in Figure 4, the results of Experiment 4 differed dramatically from those of Experiment 3. Whereas 24 out of 31 children (77%) in Experiment 3 chose the narrow latent scope coin when the coins were equally probable, only 11 out of 32 children (34%) chose the narrow latent scope coin in Experiment 4 where the wide latent scope coin was more probable. Thus, children in Experiment 4 chose the narrow latent scope explanation less often than children in Experiment 3 ($p < .001$, Fisher's exact test), and no longer preferred the narrow latent scope explanation over the wide latent scope explanation. In fact, if anything, children showed a trend to favor the *wide* latent scope explanation ($p = .11$, sign test), a preference that did not differ between 4- and 5-year-olds ($p = .46$, Fisher's exact test).

These results provide evidence that young children are able to consider information about an explanation's scope and prior probability. Like the simplicity bias (Bonawitz & Lombrozo, 2012), the latent scope bias is greatly reduced by strong prior odds. Although children in Bonawitz and Lombrozo (2012) preferred simple explanations over complex explanations when the two types of explanations had 1:1 odds, their preference for the simple explanation was significantly reduced when the odds of the simple explanation decreased to 1:6. When considered in light of our results in Experiment 4, it seems that children do not blindly rely on explanatory virtues; instead they use these explanatory virtues in concert with other sources of evidence in a flexible manner.

That said, when explanatory virtues conflict with strong prior odds, children seem to rely on probability less than one might expect. Even when children in Bonawitz and Lombrozo (2012) were presented with 1:6 odds in favor of the complex explanation, they still chose the simple explanation 30% of the time. Likewise, when children in our Experiment 4 were presented with 1:4

odds in favor of the wide latent scope explanation, they still chose the narrow latent scope explanation 34% of the time. Thus, although children clearly take probabilistic information into account, they are not narrowly focused on maximizing posterior probability. Instead, they continue to attend to explanatory virtues, even in the face of overwhelming odds.

Finally, these results help to rule out an alternative explanation of children's latent scope preference – that children chose the 1-effect coin merely because that coin corresponded to the one effect they could observe (a perceptual matching bias). Since this bias was influenced by probabilistic evidence, children must be considering multiple sources of evidence rather than blindly perceptually matching. Comparisons across experiments also speak against this possibility for two reasons: First, the effect was robust across very different tasks, whereas perceptual matching should be more fragile, depending on task context. Second, there were no age differences within any of the experiments, whereas perceptual matching should be stronger at younger ages.

General discussion

Children may be scientists, but what kind of scientists are they? Do they only seek to maximize posterior probability, or do they take account of an explanation's *loveliness* – its explanatory virtues – over and above its probability? Four experiments support the latter possibility – that children are not just little Bayesians, but little Einsteins too.

Experiments 1 and 3 demonstrated that children, like adults, show evidence of a latent scope bias, favoring explanations that make fewer unverified predictions, both when reasoning about verbal stories (Experiment 1) and a physical device (Experiment 3). Experiments 1 and 2 provided initial evidence that children's latent scope bias may be driven by the same component processes as in adults – manifest scope (tested in Experiment 1) and inferred evidence (tested in Experiment 2). Finally, Experiment 4 revealed that although this bias can lead to inferences that fail to maximize posterior probability (e.g. in Experiments 1 and 3), it can be attenuated by strong probabilistic evidence (Experiment 4). Thus, children rely on some of the same explanatory virtues as adults, attending to these cues regarding an explanation's 'loveliness' over and above its posterior probability.

This tendency to attend to explanatory virtues can lead to inferences that seem non-normative in the narrow sense, failing to maximize the posterior probability for the case at hand (Douven & Shupbach, 2015; Johnson,

Jin & Keil, 2014a; Johnson et al., 2014c, 2016; Khemlani et al., 2011; Lombrozo, 2007; Pacer et al., 2013; see van Fraassen, 1989). However, explanatory virtues such as scope, generality, and simplicity are not arbitrary – they are often good proxies for posterior probability in the sense that they pick out likely explanations under most circumstances (e.g. Lipton, 2004; McGrew, 2003). Thus, adults and children may not infer the lovely *instead of* the likely, but rather infer the lovely as a *means* of inferring the likely – a strategy that usually succeeds, but which can also lead to reasoning errors in some situations.

Where do these virtue-based explanatory inferences come from? Given children's adeptness at some forms of probabilistic reasoning (e.g. Gopnik et al., 2004; Gweon et al., 2010; Schulz & Bonawitz, 2007; Schulz et al., 2007), they may learn the explanatory virtues as short-cuts, abstracted over many episodes of probabilistic reasoning. For example, if children had many experiences deciding between simple and complex explanations, and concluded each time that the simpler explanation was more probable, they might extract the heuristic that simpler explanations are better (as suggested in Bonawitz & Lombrozo, 2012). Similarly, in the case of latent scope, if children had many pedagogical experiences in which lack of evidence signaled evidence of absence, they might extract the heuristic that unverified predictions can be inferred as false (although pedagogical inferences do not explain the adult findings; Johnson et al., 2016). Alternatively, if adult reasoning is accomplished not by direct probabilistic calculations, but instead by approximating these (computational-level) calculations using (algorithmic-level) heuristics such as explanatory virtues, then perhaps normative reasoning is instead built upon this heuristic foundation from the start.

The current results, together with those of Bonawitz and Lombrozo (2012), begin to address this question. In both sets of studies, children as young as 4 failed to maximize the posterior probability because they relied on explanatory virtues – leading to a bias against explanations making unverified predictions (in the current experiments) and a bias toward simpler explanations (Bonawitz & Lombrozo, 2012). Given the magnitude of these effects, even young children seem to have strong intuitions about explanatory virtue, in line with the idea that explanatory virtues may provide the scaffolding for probabilistic reasoning.

That said, 4-year-olds are not neonates – a great deal of learning happens in the first 4 years of life. Future research should examine both simplicity and latent scope preferences at younger ages to provide further evidence. Our suspicion is that even infants would show simplicity and latent scope preferences – if found, these results would strongly support the primacy of explanatory virtues. Given that some research already suggests that infants use explanation-based reasoning in some contexts (for a review, see Baillargeon, Li, Gertner & Wu, 2011), it would be fascinating to see whether infants incorporate adult-like explanatory virtues into their reasoning process. An equally open question concerns the evolutionary origin of these explanatory virtues. Given that non-human primates, and even rats, appear to be capable of some forms of causal reasoning (e.g. in a blicket detector task, Edwards, Rottman, Shankar, Betzler, Chituc et al., 2014; see also Blaisdell, Sawa, Leising & Waldmann, 2006), virtue-based explanatory inference may also be phylogenetically ancient.

If explanatory virtues are mechanisms for realizing probability computations, then the resulting biases may be best viewed, not as inferential failings, but as signatures of a grander method – an arsenal that may contain myriad explanatory strategies working in concert – that we can use to understand our environment, to explain what happens, and to make sense of the world in ways that are by and large adaptive from an early age. Contra Keats, beauty may not be the very essence of truth – but the explanatory virtues may suffice to get by, most of the time.

Acknowledgements

References

Agresti, A., & Coull, B.A. (1998). Approximate is better than 'exact' for interval estimation of binomial proportions. *American Statistician*, **52**, 119–126.

Baillargeon, R., Li, J., Gertner, Y., & Wu, D. (2011). How do infants reason about physical events. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development*, Vol. **2** (pp. 11–48). Oxford: Wiley-Blackwell.

Blaisdell, A.P., Sawa, K., Leising, K.J., & Waldmann, M.R. (2006). Causal reasoning in rats. *Science*, **311**, 1020–1022.

Bonawitz, E.B., & Lombrozo, T. (2012). Occam's rattle: children's use of simplicity and probability to constrain inference. *Developmental Psychology*, **48**, 1156–1164.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N.D., Spelke, E., et al. (2011). The double-edged sword of pedagogy: instruction limits spontaneous exploration and discovery. *Cognition*, **120**, 322–330.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W.J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York: Academic Press.

Cimpian, A., & Steinberg, O.D. (2014). The inherence heuristic across development: systematic differences between children's and adults' explanations for everyday facts. *Cognitive Psychology*, **75**, 130–154.

Corriveau, K.H., & Kurkul, K.E. (2014). 'Why does rain fall?': children prefer to learn from an informant who uses noncircular explanations. *Child Development*, **85** (5), 1827–1835.

Douven, I., & Schupbach, J.N. (2015). The role of explanatory considerations in updating. *Cognition*, **142**, 299–311.

Edwards, B.J., Rottman, B.M., Shankar, M., Betzler, R., Chituc, V., et al. (2014). Do capuchin monkeys (*cebus apella*) diagnose causal relations in the absence of a direct reward? *PLoS ONE*, **9** (2), e88595.

Frazier, B.N., Gelman, S.A., & Wellman, H.M. (2009). Preschoolers search for explanatory information within adult–child conversation. *Child Development*, **80** (6), 1592–1611.

Frazier, B.N., Gelman, S.A., & Wellman, H.M. (2016). Young children prefer and remember satisfying explanations. *Journal of Cognition and Development*, doi:10.1080/15248372.2015.1098649

Gopnik, A., Glymour, C., Sobel, D.M., Schulz, L.E., Kushnir, T., et al. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review*, **111**, 3–32.

Gweon, H., Tenenbaum, J.B., & Schulz, L.E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, USA*, **107**, 9066–9071.

Hussak, L.J., & Cimpian, A. (2015). An early-emerging explanatory heuristic promotes support for the status quo. *Journal of Personality and Social Psychology*, **109** (5), 739–752.

Jeffrey, R.C. (1965). *The logic of decision*. New York: McGraw-Hill.

Johnson, S.G.B., Jin, A., & Keil, F.C. (2014b). Simplicity and goodness-of-fit in explanation: the case of intuitive curve-fitting. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 701–706). Austin, TX: Cognitive Science Society.

Johnson, S.G.B., Johnston, A.M., Toig, A.E., & Keil, F.C. (2014c). Explanatory scope informs causal strength inferences. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2453–2458). Austin, TX: Cognitive Science Society.

Johnson, S.G.B., & Keil, F.C. (2014). Causal inference and the hierarchical structure of experience. *Journal of Experimental Psychology: General*, **143**, 2223–2241.

Johnson, S.G.B., Rajeev-Kumar, G., & Keil, F.C. (2014a). Inferred evidence in latent scope explanations. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 707–712). Austin, TX: Cognitive Science Society.

Johnson, S.G. B., Rajeev-Kumar, G., & Keil, F.C. (2016). Sense-making under ignorance. Manuscript in press at Cognitive Psychology.

Johnston, A.M., Sheskin, M., Johnson, S.G.B., & Keil, F.C. (2016). Preferences for explanation generality develop early in biology, but not physics. Manuscript under revision.

Khemlani, S.S., Sussman, A.B., & Oppenheimer, D.M. (2011). *Harry Potter* and the sorcerer's scope: latent scope biases in explanatory reasoning. *Memory & Cognition*, **39**, 527–535.

Legare, C.H., & Lombrozo, T. (2014). Selective effects of explanation on learning during early childhood. *Journal of Experimental Child Psychology*, **126**, 198–212.

Lipton, P. (2004). *Inference to the best explanation* (2nd edn.). London: Routledge.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, **55**, 232–257.

Lombrozo, T. (2012). Explanation and abductive inference. In K.J. Holyoak & R.G. Morrison (Eds.), *Oxford handbook of thinking and reasoning* (pp. 260–276). Oxford: Oxford University Press.

McGrew, T. (2003). Confirmation, heuristics, and explanatory reasoning. *British Journal for the Philosophy of Science*, **54**, 553–567.

Mercier, H., Bernard, S., & Clément, F. (2014). Early sensitivity to arguments: how preschoolers weight circular arguments. *Journal of Experimental Child Psychology*, **125**, 102–109.

Pacer, M., Williams, J., Chen, X., Lombrozo, T., & Griffiths, T.L. (2013). Evaluating computational models of explanation using human judgments. In A. Nicholson & P. Smyth (Eds.), *Proceedings of the 29th Conference on Uncertainty in Artificial Intelligence* (pp. 498–507). Corvallis, OR: AUAI Press.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Francisco, CA: Morgan Kaufmann.

Peirce, C.S. (1997). *Pragmatism as a principle and method of right thinking: The 1903 Harvard lectures on pragmatism* (P.A. Turrisi, Ed.). Albany, NY: State University of New York Press. (Original work published 1903.)

Read, S.J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: a parallel distributed processing account. *Journal of Personality and Social Psychology*, **65**, 429–447.

Rips, L.J. (2002). Circular reasoning. *Cognitive Science*, **26** (6), 767–795.

Robinson, E.J., Rowley, M.G., Beck, S.R., Carroll, D.J., & Apperly, I.A. (2006). Children's sensitivity to their own relative ignorance: handling of possibilities under epistemic and physical uncertainty. *Child Development*, **77**, 1642–1655.

Samarapungavan, A. (1992). Children's judgments in theory choice tasks: scientific rationality in childhood. *Cognition*, **45**, 1–32.

Schulz, L.E., & Bonawitz, E.B. (2007). Serious fun: preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, **43**, 1045–1050.

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**, 1124–1139.

Schulz, L.E., & Sommerville, J.A. (2006). God does not play dice: causal determinism and preschoolers' causal inferences. *Child Development*, **77**, 427–442.

Sussman, A.B., Khemlani, S.S., & Oppenheimer, D.M. (2014). Latent scope bias in categorization. *Journal of Experimental Social Psychology*, **52**, 1–8.

Van Fraassen, B.C. (1989). *Laws and symmetry.* Oxford: Oxford University Press.

Walker, C.M., Lombrozo, T., Legare, C.H., & Gopnik, A. (2014). Explaining prompts children to privilege inductively rich properties. *Cognition*, **133** (2), 343–357.

Walker, C.M., Williams, J.J., Lombrozo, T., & Gopnik, A. (2012). Explaining influences children's reliance on evidence and prior knowledge in causal induction. In N. Miyake, D. Peebles & R.P. Cooper (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1114–1119). Austin, TX: Cognitive Science Society.

Wellman, H.M. (2011). Reinvigorating explanations for the study of early cognitive development. *Child Development Perspectives*, **5** (1), 33–38.

Zee, A. (1999). *Fearful symmetry: The search for beauty in modern physics.* Princeton, NJ: Princeton University Press.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Data S1. Sample items for Experiments 1 and 2.