DOI: 10.1111/cdev.13647

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

Children selectively endorse speculative conjectures

Junyi Chu | **Laura E. Schulz**

MIT, Cambridge, Massachusetts, USA

Correspondence
Junyi Chu, MIT, 77 Massachusetts Ave, Cambridge, MA 02139, USA.
Email: junyichu@mit.edu

Funding information
NSF Center for Brains Minds and Machines, Grant/Award Number: 1231216

Abstract

Young children are epistemically vigilant, attending to the reliability, expertise, and confidence of their informants and the prior probability and verifiability of their claims. But the pre-eminent requirement of any hypothesis is that it provides a potential solution to the question at hand. Given questions with no known answer, the ability to selectively adopt new, unverified, speculative proposals may be critical to learning. This study explores when people might reasonably reject known facts in favor of unverified conjectures. Across four experiments, when conjectures answer questions that available facts do not, both adults (n = 48) and children (4.0–7.9 years, n = 241, of diverse race and ethnicity) prefer the conjectures, even when the conjectures are preceded by uncertainty markers or explicitly violate prior expectations.

> Bold ideas, unjustified anticipations, and speculative thought, are our only means for interpreting nature … (Popper, The Logic of Scientific Discovery)

INTRODUCTION

The history of science is full of remarkable discoveries but it is also renowned for speculative conjectures that were ultimately abandoned. Scientists posited the existence of a luminiferous aether to explain how light traveled through a vacuum, phlogiston to explain the behavior of fire, and elan vital to explain life itself. All of these ideas were ultimately discredited but today, other suppositions that have not been directly confirmed (e.g., the existence of dark matter and dark energy) play powerful explanatory roles. Faced with otherwise unanswered questions, we must sometimes entertain claims whose primary value lies, not in how certain we are of their truth, but in how certain we are that—if they were true—they would provide a solution to our problems.

But our willingness to entertain potentially explanatory but unproven speculations extends far beyond the realm of scientific inquiry. The comedian Stephen Colbert mocked politicians' readiness to accept plausible but unsubstantiated arguments as a predilection for "truthiness" over "truth" (cf: at the time of this writing, the conjecture that disinfectants might kill COVID-19 inside our bodies as well as out). However, we suggest that our ability to accept proposals based merely on the possibility that they could answer a question is not (just) a bug but also a feature of human cognition. Conjectures go beyond available knowledge and data but they are not entirely unconstrained; even the wildest speculation must, in principle if not in practice, provide an answer to the question under discussion. Insofar as one of the most challenging problems of human cognition is not distinguishing among competing hypotheses but generating them in the first place, even proposals advanced without evidence, by uncertain speakers, and in tension with prior expectations may be worth considering if they offer possible solutions to otherwise unresolved problems. Further investigation can then establish whether the hypotheses should be pursued or rejected.

We suggest that a willingness to entertain claims merely on the basis of their explanatory power is a pervasive aspect of human cognition, beginning in very early childhood. However, while many studies have looked at how children evaluate both the quality of the explanations they receive and the reliability of their informants, such work has focused almost uniformly on whether children correctly reject improbable, unreliable, or unsubstantiated claims in favor of information that is trustworthy, verified, and consistent with the integration of evidence and prior knowledge. To our knowledge, no studies have looked at situations in which either adults or children might (appropriately) reject known information in favor of the unknown. Here we look at the conditions under which people might reasonably endorse conjectural claims.

Contrary to early assumptions about the credulity of children (Clark, 1995; Piaget, 1929; Prentice & Gordon, 1987; Prentice et al., 1978), and despite the importance of pretense and fantasy in children's lives (Harris, 2000; Lillard, 2001; Lillard et al., 2011; Sharon & Woolley, 2004; Taylor et al., 1993; Walker et al., 2015; Weisberg & Gopnik, 2013; Woolley, 1995), even 3- and 4-year-olds do not adopt fanciful, speculative claims willy nilly (see Harris et al., 2018; Ronfard et al., 2018; Shafto et al., 2012; Sobel & Kushnir, 2013; Sperber et al., 2010; Woolley & Ghossainy, 2013). Children evaluate the reliability of their informants in increasingly sophisticated ways from preschool through middle childhood (Clément, 2010; Gweon et al., 2014; Koenig et al., 2004; Koenig & Harris, 2005; Koenig & Woodward, 2010; Pasquini et al., 2007). They are sensitive to the strength of the evidence they observe (Bridgers et al., 2016; Butler et al., 2018), the prior probability of testimony (Chan & Tardif, 2013; Clément et al., 2004; Jaswal, 2004; Koenig et al., 2004; Koenig & Echols, 2003; Ma & Ganea, 2010), the informant's past accuracy, knowledge, and expertise (Danovitch & Keil, 2004; Koenig & Harris, 2005; Koenig & Jaswal, 2011; Kushnir, 2013; Kushnir et al., 2013; Landrum et al., 2013; Nguyen, 2012; Sobel & Corriveau, 2010; Sobel & Macris, 2013; Vanderborght & Jaswal, 2009), and the situational and epistemic constraints the informant is under (Butler et al., 2018; Einav & Robinson, 2011; Flavell, 1988; Nurmsoo & Robinson, 2009; Senju et al., 2011).

This does not mean that all of children's evaluation of informants is epistemically justified. Preschoolers have a general bias in favor of agents who are friendly (Brosseau-Liard & Birch, 2010; Landrum et al., 2013), familiar (Reyes-Jaquez & Echols, 2013), attractive (Bascandziev & Harris, 2014; Fusaro et al., 2011), members of their in-group (Elashi & Mills, 2014; Kinzler et al., 2011; MacDonald et al., 2013; Plötner et al., 2015; Wood et al., 2013), or part of a majority (Chen et al., 2012; Corriveau & Harris, 2010; DiYanni et al., 2015; Morgan et al., 2014). Children are also influenced by the confidence with which informants assert their claims. Preschoolers are more likely to endorse novel explanations and labels advanced with confidence than those provided by a speaker who is hesitant or expresses uncertainty (Jaswal & Malone, 2007; Kominsky et al., 2016; Sabbagh & Baldwin, 2001; Tenney et al., 2011). However, even these prima facie non-epistemic biases may be reasonable routes to learning insofar as friendly, familiar, in-group members who are backed by a majority consensus may typically also be the most likely sources of reliable information.

Collectively, this literature suggests that children might be very likely to reject information that is unverified or unverifiable, especially if the speaker conveys uncertainty or the information is itself unlikely or unexpected. However, children are also interested in getting answers to their questions. Children are notorious for asking questions themselves (as many as 76 an hour; Chouinard, 2007), and although some are requests for permission, or for information redundant with facts the child already knows (Legare et al., 2013; Ruggeri et al., 2016), many are requests for novel information and explanations (Callanan & Oakes, 1992; Chouinard, 2007; Frazier et al., 2009; Mills et al., 2010, 2011).

Preschoolers also understand many structural aspects of explanation and can evaluate respondents' answers on those bases alone. If a respondent simply restates a child's question, asserts norms, re-describes events, or reacts personally instead of responding to the query, the child is likely to repeat the question (Chouinard, 2007; Frazier et al., 2009; Kurkul & Corriveau, 2017; Tizard & Hughes, 1984). Preschoolers favor claims supported by strong arguments over circular ones, and circular arguments over unsupported opinions (Corriveau & Kurkul, 2014; Mercier et al., 2014; Mills et al., 2017) and evaluate explanations based on how many observations an explanation accounts for, how simple and internally coherent it is, and how probable it is given observed data and their prior knowledge (Bonawitz & Lombrozo, 2012; Johnston et al., 2016; Lombrozo, 2011; Walker et al., 2017).

Children can also use data-independent criteria to evaluate hypotheses that lack direct evidential support. In addition to favoring explanations that are simple, broad and coherent (Bonawitz & Lombrozo, 2012), children can use properties of the explanandum when choosing among equally probable hypotheses. When asked to match observed events to their probable causes, preschoolers expected discrete and continuous affordances to control discrete and continuous phenomena, respectively, without observing any covariation data (Magid et al., 2015; Tsividis et al., 2015). This suggests that children might be sensitive to abstract features of causes and effects and use these features to constrain their generation and evaluation of candidate causes. Independent of the content of the domain, and in the absence of any distinguishing evidence, children might be able to use properties of the question under discussion to decide what makes for a good answer.

In the current studies, we look at whether children flexibly evaluate facts and conjectures given questions that can or cannot be answered by available information. In Experiment 1, we introduce children to short stories involving novel characters and events. We ask children to choose between factual and conjectural explanations, for questions that can or cannot be answered by information in the story. We used nonsense characters and stories in order to control for effects of prior knowledge, and we matched the answers on the degree to which they repeated the words from the question in the answer. Thus, children must consider the degree to which each response provides a potential answer to the question at hand. In Experiment 1a, we tested a wide age-range (4- to 8-year-olds) given that it was not clear to what extent children at different ages might privilege abstract features of explanations over established facts.

If children always prefer the most certain and reliable information, they should always choose the facts; if they always prefer more speculative, inventive answers, they should always choose the conjectures. However, we predict that the 6- to 8-year-olds, and possibly even the preschoolers, would prefer the known facts for questions that can be answered by information in the story and prefer conjectures for questions that cannot.

EXPERIMENT 1

We ran both an initial exploratory study (Experiment 1a) and a replication with just the 4- and 5-year-old's (Experiment 1b). Hypotheses were prespecified ahead of data collection, but not formally preregistered. For the initial experiment, we estimated a moderately large effect size in choosing facts for questions with available answers and conjectures for questions with unknown answers. We aimed to recruit 64 participants in Experiment 1a which would yield 80% power to detect an odds ratio of 5.23 (pilot testing had suggested an odds ratio of 6.93). In Experiment 1b, we tested only a younger age group, and recruited 32 participants to match the number of younger children in Experiment 1a.

METHODS

Participants

All children in this and the following experiments were recruited from an urban children's museum between January 2018 and November 2019 in the United States. Parents provided informed consent, and children received stickers for their participation. Although most of the children were white and middle class, a range of ethnicities and socioeconomic backgrounds are represented in museum attendees overall (47% European American, 24% African American, 9% Asian, 17% Latino, 4% two or more races; 29% of museum attendees visit on days when there is free or discounted admission).

In Experiment 1a, we tested 66 children, ages four to eight (M = 6.04 years, range: 4–7.93). Seven additional children participated but were excluded for either responding inaccurately on a practice question ($N = 5$), not speaking English as their primary language ($N = 1$) or for incomplete participation ($N = 1$).

In Experiment 1b, we tested 32 four- and five-year-olds (M = 5.03 years, range: 4.15–5.92). Thirteen additional children did not pass the inclusion criteria (9 failed practice; 2 did not speak English as their primary language; 1 withdrew; 1 did not respond to test questions). The exclusion rates for 4- and 5-year-olds are relatively high but can be explained almost entirely by children choosing Elmo as the correct puppet on both practice trials.

Materials and procedure

Each trial began with an illustrated story presented via three animated slides on a laptop computer. See Figure 1 for an example story. Two puppets (Elmo and Cookie Monster) were also used; the puppets sat on either side of the computer and "watched" the stories with the child. The puppets' answers were delivered by pre-recorded audio clips to avoid inadvertently biasing the children with prosodic cues.

Participants completed two training trials and four test trials. The training trials were designed to ensure that participants were paying attention and understood the task. Participants who failed the training trials were excluded from analysis and replaced. These stories depicted human characters performing common activities (i.e., riding a bike; eating ice cream) embedded in a simple narrative. The training questions could always be answered using information from the story. Each puppet provided a correct answer on one trial and an incorrect trial on the other (order counterbalanced).

Children were tested individually in a quiet room. The experimenter began by introducing participants to the computer display and the puppets (Elmo on the left and Cookie Monster on the right). The experimenter explained the task, saying: "Every time I tell you a story, I need you to remember what happened because I'm going to ask a question at the end. Elmo and Cookie Monster will tell us their answers and your job is to choose who had the better answer." On every trial, the experimenter first narrated the story and presented her question ("My question is, …?"). She then directed the question at one puppet (e.g., "Elmo, can you tell us, …?"), played its pre-recorded answer, and repeated the answer (e.g., "Elmo said because …"). The experimenter then repeated the question-answer sequence with the other puppet before repeating the question and inviting the child to make a choice ("My question was … Who do you think had the better answer for [question]?"). Positive feedback was given on the training trials ("That's right, Elmo had the better answer this time.") and neutral, encouraging feedback was given on the test trials ("Alright, let's see what's next"). Only children who correctly answered both practice questions continued to complete the four test trials.

[Slide 1] Here are some juggling Gazzers. A clown named Bozo taught them to juggle.

[Slide 2] Juggling Gazzers love to eat bananas. But the bananas all grow at the top of very tall trees and the Gazzers can't climb trees.

[Slide 3] But here the Gazzers are! Eating bananas.

In-Story Question: How did the Gazzers learn to juggle?

Out-of-Story Question: How did the Gazzers get the bananas?

Fact: Because Bozo the clown taught the banana eating Gazzers how to juggle

Conjecture: Because the Gazzers threw their balls up into the trees and knocked down the bananas.

FIGURE 1 Example of a test trial used in Experiments 1 and 2

The test trials involved imaginary creatures engaging in different activities (making a hat, sneezing from allergies, dropping a toy down a deep hole, juggling). Two question-answer pairs were used on each story: One question could not be answered with the conjecture offered but could be answered with a Fact mentioned in the story (In-Story Question); the other question could not be answered with any facts in the story and could only be answered with a Conjecture (Out-of-Story question). Regardless of question type, Elmo always provided the Fact answer and Cookie Monster always provided the Conjecture answer. Elmo always provided his answer first. To cover a range of explanatory question types, test trials included both "why" and "how" questions. In Experiment 1a, two "how" questions came first; in Experiment 1b, two "why" questions came first. In both experiments, we counterbalanced two between-participant factors: (1) item order (whether the first test trial was an In-Story or Out-of-Story question) and (2) story-question match, resulting in four-story sequences. Thus, while all participants heard all stories and answers, half the participants heard any given story presented with a question that could only be answered with a fact and half heard the story presented with a question that could be answered only with a conjecture (see Table 1).

Results

In Experiment 1a, our primary research question was whether participants would choose the appropriate explanation on each trial: Facts for In-Story questions and Conjectures for Out-of-Story questions. Figure 2 shows children's responses by question type. Across all age groups and conditions, children successfully matched answers with question types (3.17 of 4 trials; SD = 0.71; $t(65) = 13.27$, $p < .001$). A third of the children (23/66) chose the appropriate answer at ceiling (binomial $p < .001$).

We looked at whether children's responses varied by age and question type using a logistic mixed-effects model. This model predicted children's response (0 = fact, 1 = conjecture) from age (in months, mean-centered), question type (0 = In-Story, 1 = Out-of-Story), and an interaction of age and question type, with random intercepts for subject and story. There was no effect of age ($p = .080$ by asymptotic Wald test) or an interaction of age and question type ($p = .155$). As predicted, there was a main effect of question type ($\beta = 3.23$; OR = 25.24, 95% CI [11.9, 53.7]; $p < .001$). Children endorsed the fact more often on In-Story questions (M = 1.76 of 2 trials, SD = 0.43) and the conjecture more often on Out-of-Story questions (M = 1.41 of 2 trials, SD = 0.58).

Given that there was no effect of age in Experiment 1a, we looked at whether the same results would hold looking only at the 4- and 5-year-olds. Experiment 1b was identical to Experiment 1a, except that, as noted above, we presented "how" stories before the "why" stories (see Supporting Information for post hoc exploratory analyses of performance by "how" and "why"). Four- and 5-year-olds successfully matched answers with question types (M = 2.84 of 4 trials, SD = 1.14; $t(31) = 4.19$, $p < .001$). As in Experiment 1a, approximately one-third of children (12/32 or 38%) chose the appropriate explanation at ceiling ($p < .001$).

TABLE 1 Questions and candidate answers used in Experiments 1 and 2

Story	In-story question	Out-of-story question	Fact answer	Conjecture answer
Training 1	How did Tommy get to the castle?	—	He rode a bike to the castle	He walked to the castle
Training 2	What is Tommy's favorite ice cream?	—	Chocolate	Strawberry
Test 1	Why are the Wugs sneezing?	Why are the Feps furry?	Because the Wugs are allergic to the Feps' fur	Because the Feps go up to the mountains and the fur keeps the Feps warm
Test 2	Why are the small Daxes wearing hats?	Why are the hat-making Blickets bigger than the Daxes?	Because the big Blickets made the hats for them	Because the Blickets are older than the Daxes
Test 3	How did the banana eating Gazzers learn to juggle?	How did the Gazzers get the bananas?	Because Bozo the clown taught the banana eating Gazzers how to juggle	Because the Gazzers threw their balls up into the trees and knocked down the bananas
Test 4	How did the Duff's toy fall into the deep hole?	How did the Duffs rescue their toy?	Because the Duffs' hair was in their eyes and they couldn't see, and they tripped and dropped their toy	Because the Duffs tied their long hair into a rope and made a ladder with it and used it to climb down and get the toy

Note: Participants were excluded for responding incorrectly to either Training trial.

Next, we fit a logistic mixed-effects model to predict children's choices from age, question type, and age by question type interaction, with random intercepts for subject and story. Replicating Experiment 1a, we found a main effect of question type ($\beta = 2.22$; OR = 9.24, 95% CI [3.4, 25.4]; $p < .001$); children endorsed the Fact more often on In-Story questions (M = 1.44 of 2 trials, SD = 0.84) and the Conjecture more often on Out-of-Story questions (M = 1.41 of 2 trials, SD = 0.71). There was no effect of age ($p = .281$) or age by question type interaction ($p = .181$).

Discussion

The results of Experiments 1a and 1b suggest that children as young as four and five flexibly consider the explanatory demands of the question under discussion. Children did not show a consistent preference either for known facts or novel information. Rather, children appropriately used the question to guide their evaluation of possible answers. When questions could be answered by available facts, children preferred factual answers; when they could not, children rejected the established facts in favor of conjectural claims for which they had no independent evidence.

EXPERIMENT 2

Experiments 1a and 1b showed that 4- and 5-year-olds were willing to answer questions with novel unverified conjectures rather than known facts; however, the forced choice design meant we cannot tell whether children actively endorsed conjectures for otherwise unanswered questions or whether they simply rejected facts that failed to answer questions satisfactorily. In Experiment 2, we ask children to rate each response independently and manipulate explanation type as a between-subjects comparison such that children never got to compare facts against conjectures.

Also, in the preceding experiments, we did not give children any direct information about the empirical status of the conjectures; it is possible that the children in Experiment 1 may have accepted the conjectures because they failed to recognize that they were indeed speculative and unverified. In Experiment 2, we add an Uncertain Conjectures condition where we emphasize the speculative nature of the conjectures by prefacing them with explicit uncertainty markers (*I don't know, but maybe …*). Abundant evidence suggests that 4- and 5-year-olds preferentially endorse claims from speakers who are knowledgeable and confident over those from speakers who admit ignorance or uncertainty (saying *I don't know*, *Hmm*, or *maybe*; Jaswal & Malone, 2007; Moore et al., 1989; Sabbagh & Baldwin, 2001; Sabbagh & Shafman, 2009; Tenney et al., 2011). If in independent judgments,

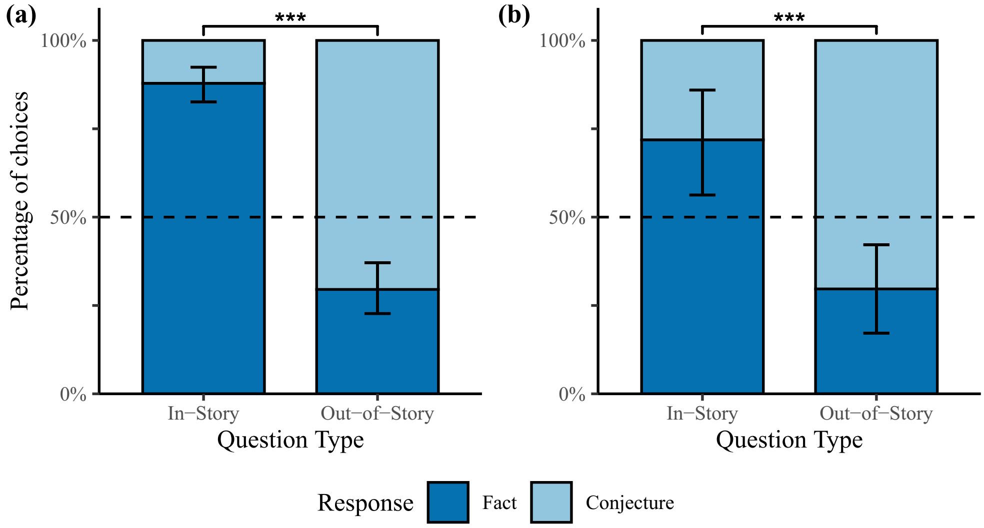

FIGURE 2 Children's ratings (averaged across two test trials for each question type) in Experiments 1a (panel a: $N = 66$, mean: 6.04 years; range: 4.00–7.93) and 1b (panel b: $N = 32$; mean: 5.03 years; range: 4.15–5.92). Children were more likely to choose facts when the question could be answered by information in the story and conjectures when it could not. Error bars show bootstrapped 95% confidence intervals. Paired *t*-test, $^{***}p < .001$

children appropriately endorse conjectures even when they are advanced by uncertain speakers, this would be a strong evidence that children value conjectures based simply on their ability to answer otherwise unresolved questions.

Note that if children succeed in this task, this would not be the first study to show that children sometimes prefer hesitant speakers to confident ones. Indeed, children show precisely this preference when a hesitant speaker is appropriately calibrated to her uncertainty (e.g., because she lacks epistemic access) and a confident (but ignorant) speaker is mis-calibrated (Birch et al., 2020; see also; Huh et al., 2019). Critically however, there are a number of methodological differences between our task and previous work showing that children prefer informants who appropriately mark their uncertainty. Prior studies involved agents who did or did not know specific facts (e.g., the contents of a box, the name of an object; Brosseau-Liard et al., 2014; Tenney et al., 2011). Here by contrast, informants are probed for explanations of causal events. We believe children might tolerate causal conjectures without explicit uncertainty markers precisely because children may recognize that such answers *are* speculative. When it is in common ground between the child and the informant that the relevant facts are not available, it might be less important that informants convey their uncertainty explicitly. Thus, consistent with the calibration literature, we expect that children will be un-swayed by confident statements that fail to answer a question and that children will endorse conjectures that answer questions when the informant expresses uncertainty. However, insofar as children recognize conjectures as such, we predict that they will not penalize informants who fail to convey uncertainty.

We treated this as a confirmatory study and preregistered all analyses and predictions on the Open Science Framework (https://osf.io/zpq3r). Power analysis using simulations from pilot data indicated that a sample of 32 participants per condition would yield 80% power to detect a moderate interaction of Explanation condition by Question type.

Methods

Participants

Participants were 95 four- and five-year-olds ($M = 5.03$, range: 4.03–5.99) recruited and tested as in Experiment 1. Thirty-four additional children participated but were excluded for responding inaccurately on an inclusion trial ($N = 26$), not speaking English as a native language ($N = 2$), experimenter error ($N = 5$) or failing to complete the study ($N = 1$). Note that the exclusion rate (21% of 121 initial participants) is high and similar to Experiment 1b. In this case, excluded participants overwhelmingly (21 of 26 children) correctly put Elmo in the “good cup” on the first trial and then also (incorrectly) put Cookie Monster in the “good cup”. That is, in a forced choice of Elmo or Cookie Monster (Experiment 1b) children showed an “Elmo bias”; in independent judgments (and perhaps unsurprisingly given the status of these characters in children's lives) children showed a positive response bias to both puppets. Participants were randomly assigned to the Fact, Conjecture, or Uncertain Conjecture condition and we found no condition differences in age ($M = 5.07$, 4.98, and 5.02 years, respectively; $p > .8$).

Materials and procedure

The Materials and Procedure were identical to those in Experiment 1 except as follows. Instead of using the same pair of puppets on every trial, six different cartoon characters were used, one for each trial. The characters were taken from the Muppets and each character was printed on laminated paper and glued to a wooden stick; puppets were ~15 cm tall. Each puppet appeared just once so that children could evaluate each question and answer pair independently across trials. We also used two identical blue cups (~20 tall), one labeled with a smiley face sticker (in which the child could put puppets who gave "good answers") and one left blank (for placing puppets who gave "not so good answers"). These were kept in a fixed position with the "good answer" cup on the child's left and "not so good answer" cup on the child's right. See Figure 3.

Children were told that they would hear some stories and then hear some questions and that the puppets would try to answer those questions. The experimenter explained, "Some puppets will give good answers and some puppets will give not so good answers." Children were introduced to the cup for good answers and the cup for not so good answers and asked to point to each. All participants correctly identified the two cups before proceeding to the training trials.

On each trial, the experimenter narrated a story accompanied by an animated slide deck and posed a question at the end. The children were introduced to just one puppet on each trial and the puppet responded with a pre-recorded answer (activated by the experimenter). The experimenter then asked the child, "Was that a good or not so good answer?" Children rated the puppet's response by placing them into one of two cups.

As in Experiment 1, we designed the training trials to familiarize participants with the question and explanation evaluation process, and to elicit both ratings. Training trials were the same for every child: The puppet on the first trial provided a good answer to the question; the puppet on the second trial provided a not so good answer. Children received feedback on these items to reinforce the two-cup rating system. Any child who responded incorrectly on either training trial was excluded from further analysis and replaced. Next, the experimenter presented the four test trials. As in Experiment 1, two trials involved In-Story questions and two trials involved Out-of-Story questions.

In the Fact condition, the puppet on each of the four trials responded with a verified true fact from the story (i.e., regardless of whether they were asked an In-Story or Out-of-Story question). Thus, on the In-Story trials, the puppets provided good answers and on the Out-of-Story trials, the puppets provided not so good answers. In both the Conjecture and Uncertain Conjecture condition, the puppets on each of the four trials responded with an unverified conjecture; thus, in both conditions, the puppets on the In-Story trials provided not so good answers and the puppets on the Out-of-Story trials provided good answers. In the Uncertain Conjecture condition, the answer was preceded by "I don't know, but maybe …". Test trials were counterbalanced as in Experiment 1, with eight versions per condition.

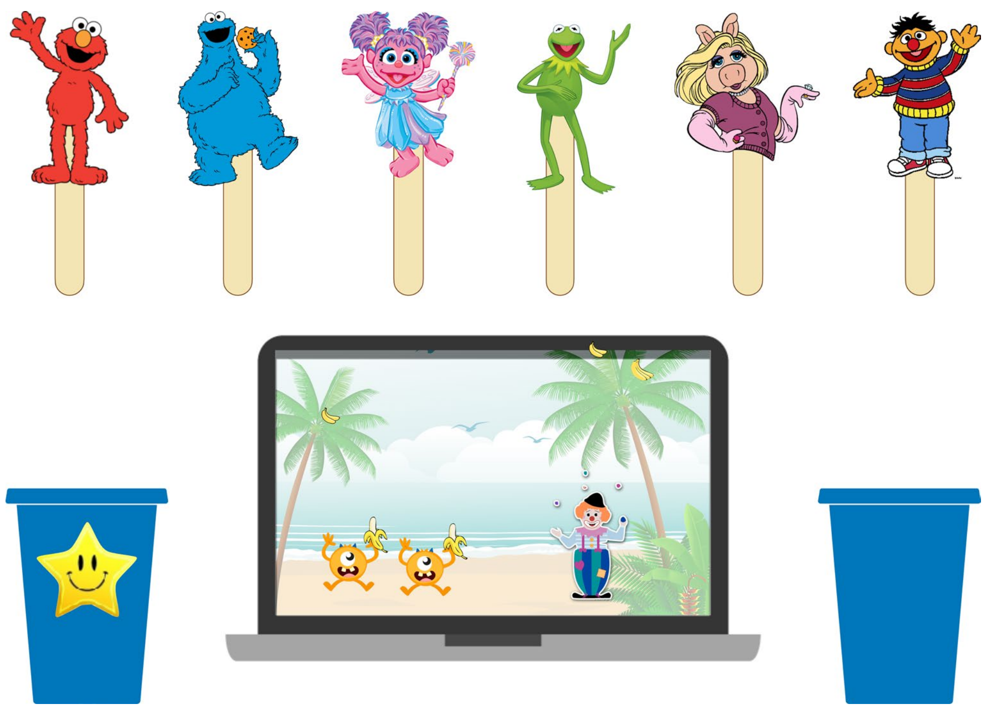

FIGURE 3 Puppets and rating cups used in Experiments 2–4

Results

Participants' responses are shown in Figure 4. As was evident in the inclusion trials children showed a positivity bias toward both puppets and rarely placed any of the puppets in the "not so good" cup. Across all conditions, children successfully matched answers with question types (M = 2.38 of 4 trials; SD = 0.92; $t(94) = 3.99$, $p < .001$), although at a lower rate than in Experiment 1. About 12% of the children (11/95) chose the appropriate answer at ceiling, not significantly different from chance (binomial $p = .051$).

Our first question was whether children would give Facts higher ratings for In-Story questions and Conjectures higher ratings for Out-of-Story questions. We predicted that the likelihood of endorsing each explanation would depend on an interaction between Explanation Condition (reference category = Fact) and Question Type (reference category = In-Story). To test this prediction, we used a mixed effects logistic regression to predict children's endorsement on each trial, including fixed effects of Explanation Condition, Question Type, and their interaction, as well as random intercepts for subject and story (model syntax: RatedAsGood ~ Explanation Condition * Question Type + (1|Subject) + (1|Story)).

As predicted, the Explanation Condition by Question Type interaction explained significant variance ($\chi^2(2) = 14.9$; $p < .001$). We conducted follow-up contrasts using estimated marginal means, with Bonferroni corrections for multiple comparisons. These contrasts found that children in the Fact condition were more likely to give positive ratings on In-Story questions (84%) than Out-of-Story questions (68%), although the result did not reach the threshold for statistical significance ($p = .18$). In contrast, children hearing Conjectures were more likely to give positive ratings on the Out-of-Story questions (83%) than the In-Story questions (61%; $p = .04$). Children hearing Uncertain Conjectures were also more likely to give positive ratings on Out-of-Story questions (76%) than the In-Story questions (52%; $p = .05$).

Our second question was whether children's ratings of conjectures would be affected by expressions of uncertainty. Follow-up contrasts comparing children's ratings in the Conjecture and Uncertain Conjecture conditions found no difference for either In-Story questions or Out-of-Story questions (zs < 1). We also tested for any interaction of question type and condition type by repeating the previous regression analysis including just the Conjecture and Uncertain Conjecture conditions. This regression analysis did not find a significant Explanation Condition by Question Type interaction ($z = -0.06$, $p = .954$) or a main effect of Explanation Condition ($z = -0.83$, $p = .404$). However, there was a main effect of Question Type ($\beta = 1.23$, OR = 3.43, 95% CI [1.40, 8.41], $z = 2.69$, $p = .007$), showing that children consistently rated conjectures more highly for Out-of-Story questions than In-Story questions. Thus, children's judgments were not significantly impacted by explicit expressions of uncertainty.

Finally, we asked whether children's sensitivity to the match between question type and explanation type was driven by an active recognition of appropriate answers, a rejection of inappropriate answers, or both. We used one-sided Wilcoxon signed rank tests to compare children's ratings against chance for each of the six combinations of Explanation Condition by Question Type (i.e., each bar in Figure 4). Correcting for multiple comparisons, we found that children rated explanations "good" significantly more often than chance when explanations were appropriate to the question type (e.g., Facts for In-Story questions and Conjectures for out-of-Story questions, ps < .0083). However, when explanations were inappropriate to the question type, children did not reject these explanations more often than chance (ps > .0083).

Discussion

Although children in Experiment 2 were inclined to endorse all the answers they were given, they were nonetheless sensitive to the relationship between questions and answer types. Replicating Experiment 1, children in Experiment 2 preferentially endorsed facts for questions whose answer could be found in the story and conjectures for questions whose answer was unknown. Critically in this context, adding explicit markers of ignorance and uncertainty ("I don't know, but maybe …") did not impact children's endorsement of conjectural explanations.

These results are compatible with a growing literature showing that children use situational constraints to evaluate testimony; that is, children assess whether speakers' claims are justified given their epistemic access (e.g., Birch et al., 2020; Brosseau-Liard et al., 2014; Huh et al., 2019; Tenney et al., 2011; see Koenig et al., 2019 for review). Consistent with this literature, we find that 4- and 5-year-olds readily accept hesitant speakers who offer speculations as answers to unresolved questions. However, in contrast to previous work suggesting that children penalize overconfident speakers who fail to convey hesitancy when reporting uncertain information, children here were happy to endorse speakers who advanced conjectures without uncertainty markers. As noted, we believe that this is because the uncertainty marker might be redundant in these contexts insofar as both the children and informant recognize that the answers were indeed speculative.

One limitation in interpreting the results of Experiment 2 is that children were generally inclined to endorse all the answers that they were given. Future research might use a more sensitive measure of children's judgments, such as a rating scale with more than two response options, or by

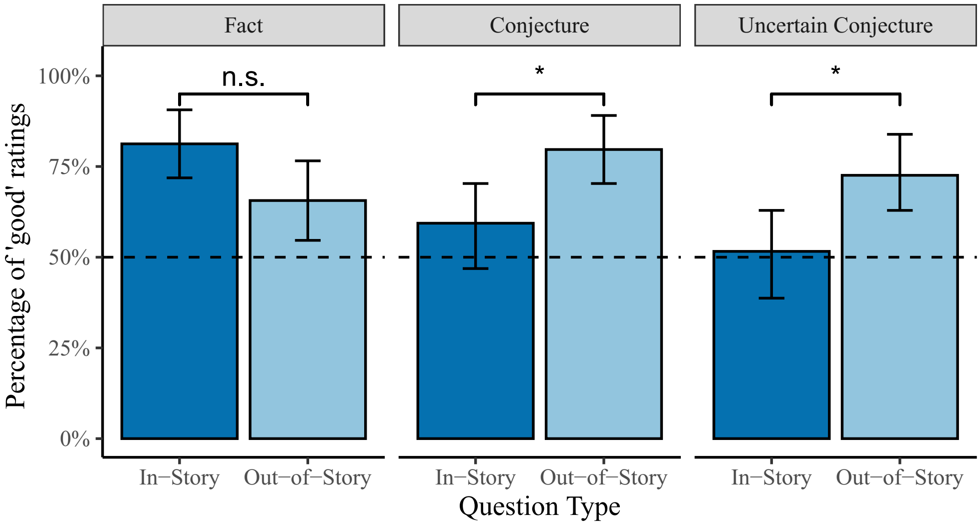

FIGURE 4 Children's ($N = 95$; mean: 5.03 years; range: 4.03–5.99) ratings in Experiment 2 averaged across two test trials for each question type. When given only factual answers, children tended to endorse these answers across the board. In contrast, children were more likely to endorse conjectural answers (whether offered neutrally or with an explicit uncertainty marker) only when the question could not be answered by information in the story. Error bars show bootstrapped 95% confidence intervals. (Paired t-tests, $*p < .05$)

asking children to explain their ratings. For consistency and ease of comparison, however, we used the same binary rating scale in Experiments 3–4.

EXPERIMENT 3

In Experiment 2, children endorsed conjectures that contained information not substantiated by the stories. However, although the information was novel, it was not especially surprising. Would children be willing to endorse conjectures that answered otherwise unanswered questions if the conjectures were improbable given the children's prior expectations?

Abundant research has testified to young children's ability to integrate evidence with prior knowledge to draw rational inferences (Gopnik & Wellman, 2012; Legare et al., 2010; Schulz, 2012a; Schulz et al., 2007; Sobel & Kushnir, 2013; Tenenbaum et al., 2011; Xu & Tenenbaum, 2007). Children's sensitivity to both the prior probability of hypotheses and the data in their favor might reasonably lead children to reject conjectures that are supported by neither.

However, a remarkable feature of human learning is that we can and do go beyond both the data and current knowledge to advance new, and even initially prima facie unlikely ideas. After the explosion on Apollo 13, the astronauts and ground crew had to improvise a way of connecting canisters for removing carbon dioxide to the lunar module; they succeeded using cardboard ripped from their training manual, towels, and duct tape. The proposal was endorsed, not because of the weight of evidence in its favor, nor because cardboard, towels, and duct tapes were typically used for these ends, but because—faced with a problem and no apparent solution—a speculative proposal that might solve the problem could be valued on those grounds alone. Obviously, Apollo 13 was an extraordinary event; in everyday cognition, our willingness to endorse otherwise unfounded conjectures may be more likely to inflame superstitions or perpetuate conspiracy theories than save lives. Still, the ability to value hypotheses simply because they could answer questions or solve problems (an ability that we, in homage to William James, will refer to as cognitive pragmatism) may contribute to human learners' distinctively powerful ability to generate new knowledge about the world.

In Experiment 3, we look at participants' willingness to accept conjectures that answer the question at hand when the conjectures contradict expectations set up by the story or when the conjectures are also rare, low probability events in themselves. That is, rather than pitting conjectural answers against facts, we compare more and less plausible conjectures, in contexts where they either do or do not provide answers to the question at hand. We test both adults and a relatively wide age-range of children (as in Experiment 1a), since to our knowledge, no studies have looked at whether participants at any age will prioritize the pragmatic goal of answering questions over other considerations and endorse unlikely conjectures when they provide a potential resolution to otherwise unresolved queries. An a priori power analysis was conducted using G*Power3 to test main and interaction effects in a 2 × 2 repeated measures ANOVA, using an F-test with a medium effect size (f = .25) and an alpha of .05. Results indicated that a sample of 24 participants was required to achieve a power of 0.8.

TABLE 2 The two training trials and one of six test trials in Experiment 3

Item	Text
Training Story 1	This is Tina. Tina is having breakfast. She's eating pancakes. After breakfast, she went outside to play. When she was done, it was time for lunch. Tina comes back into the kitchen. Her brother comes into the room and asks, "Hey Tina, what did you have for breakfast?"
Question	What did Tina have for breakfast?
Likely Answer (Fact)	Tina ate some pancakes for breakfast.
Likely Non-Answer	Tina played in the tree-house for breakfast.
Training Story 2	This is Tommy. Tommy went to the ice cream shop. He bought his favorite ice cream and ate it all up! (only an empty cone was shown)
Question	What did Tommy get at the ice cream shop?
Unlikely Non-Answer	Tommy got tomato soup.
Likely Answer	Tommy got chocolate ice cream.
Test Story	This is Sally. Sally was looking forward to her best friend's birthday party. Her best friend had just mailed out the invitations, and Sally was hoping to get one soon. Sally walked to her mailbox and saw a shiny white envelope. Sally opened the envelope and jumped up and down excitedly when she read it!
Question	Why was Sally so excited?
Likely Answer	Because she got invited to her best friend's birthday party. She was so excited that she couldn't stop jumping up and down.
Unlikely Answer	Not because it was a birthday invitation, but because it was a letter from school saying she won the story competition.
Likely Non-Answer	Not because it was a birthday invitation, but because it was a notice from the library saying she forgot to return her books.
Unlikely Non-Answer	Not because it was a birthday invitation, but because it was a note from her teacher saying that she had to do extra work after school.

Note: See Supporting Information for full stimuli. Participants were excluded for incorrectly responding to the training trials. In Experiment 4, we used the same training trials and test stories and questions, but only presented the Unlikely Answers and the two Non-Answers. We also modified the conjectures in Experiment 4 to the form "It was not [a birthday invitation/…]. Sally was excited because [it was a letter from school saying she won the story competition/…]."

Methods

Participants

Twenty-four adults were recruited and tested via Amazon Mechanical Turk and paid $1.00 for participating. Ten additional adults were excluded for failing to complete the experiment (n = 2) or failing to distinguish good and bad responses on the two inclusion trials by at least a 10-point spread (n = 8). Twenty-four 4- to 7-year-olds (M = 5.95 years, range = 4.37–6.71) were recruited and tested as in the preceding experiments. Fourteen additional children participated but were excluded for responding inaccurately on the two inclusion trials (N = 7), being distracted during one or more trials (N = 1) or failing to complete the study (N = 6).

Materials

We created six test stories, each paired with one question and four candidate answers. Twenty-four unique puppets were used to present the candidate answers, with characters from the Muppets and Hey Arthur! shows and constructed like those in Experiment 2. The two "good answer" and "not so good answer" cups used in Experiment 2 were also used here.

In contrast to the previous studies where we wanted to control for prior knowledge, in Experiment 3 we wanted to leverage participants' background knowledge to evoke strong expectations about plausibility. Thus, rather than using novel characters, these stories involved human children in everyday activities. The stories always ended with the protagonist in a salient emotional or behavioral state and the questions all asked why the character was in that state. The story context always set up one answer as the most plausible answer (i.e., the "Likely Answer"). However, in contrast to the previous studies, no question was directly answered by the information in the story (i.e., we only asked Out-of-Story questions). All the candidate explanations were conjectures; they varied in whether or not they answered the question at hand and how plausible they were. That is, on every trial, the four candidate answers always crossed two factors: Whether it would answer the question if it were true (Answer/Non-answer) and how probable the conjectured event was (Likely/Unlikely). To ensure that children recognized which candidate answers violated expectations, all answers (except for the Likely Answer) began by explicitly denying the likely answer (i.e., they took the form

"Not because of … but because of …"). See Table 2 for the complete text of one trial; full stimuli are presented in Supporting Information.

As in preceding experiments, we included two training trials to familiarize participants with evaluating the answers to questions about the stories. Although we had used only fact questions in the training for Experiments 1 and 2; in the training for Experiment 3, we included a conjectural question because all of the test trials involved contrasts between conjectures. On the first training trial, we asked a factual question that could be answered by recalling information from the story. Participants rated two answers: One was true with respect to story and answered the question; the other was true but did not answer the question. On the second training trial, we asked a conjectural question that could not be answered given available information. Participants rated two explanations: One was a likely answer given the story; the other was unlikely. (Note therefore that if anything, the training trial should make participants less likely to endorse unlikely conjectures.) These training trials were the same for all participants and were used as inclusion criteria. To ensure that adults on M-Turk were following the task instructions, we used a fairly conservative inclusion criteria: Adults had to distinguish the good and bad questions by a 10 point spread in ratings; adult participants were excluded from further analysis if they failed to make this distinction (and were thus possibly responding at random) or if they reversed the ratings. Children were excluded from further analysis and replaced if they answered either training trial incorrectly.

Procedure

We used a within-participants design: participants saw all six test stories and rated all four answers for each story. The stories were presented in a fixed sequence; answers were randomized for adults and presented in pseudorandom order for children.

Adult participants used a linear rating scale ranging from Not Satisfying (0) to Very Satisfying (100) and saw all four conjectures presented on the same page. Adults did not receive any feedback on the training trials but adults were excluded from further analysis and replaced if they did not rate the better explanation at least 10 points higher than the alternative on the training trials.

Children were tested individually in a quiet room with a laptop computer. The experimenter provided feedback on training trials to reinforce the two-cup rating system. Any child who responded inappropriately on a training trial was excluded from further analysis and replaced. The procedure and rating system was identical to the one used in Experiment 2 except as follows. The experimenter read each story out loud and then asked the target question. Four puppets took turns giving each of the candidate responses (as in the previous studies, all responses were pre-recorded to avoid differential prosodic cues across trials). Children rated each conjecture as either a "good" or a "not so good" answer by placing the puppet into the appropriate cup. Children had to rate each puppet's answer before hearing the next puppet's answer.

Results

Our main question was whether participants would evaluate conjectures based not just on how plausible the conjecture was but also on how well it would answer the question, if true. We used mixed effects regression predicting responses to each conjecture. This analysis included fixed effects of Answer Type (Answer/Non-Answer) and Likelihood (Likely/Unlikely), as well as random effects for subject and story (Model syntax: Response ~ Answer Type * Likelihood + (1|Subject) + (1|Story)). Recall that adult participants provided a continuous rating (0–100) but child participants provided a binary rating (0 or 1). Thus, we used linear regressions for adults and logistic regressions for children.

Among adult participants, there was a significant effect of Answer Type ($\beta = 33.41$, 95% CI [27.9, 38.9], $z = 11.89$, $p < .001$), with higher ratings for Answers (M = 64.9, SD = 14.4) than Non-Answers (M = 10.6, SD = 19.8). The effect of Likelihood was not significant ($z = 0.05$, $p = .961$), however, there was a significant Answer Type by Likelihood interaction ($\beta = 41.73$, 95% CI [33.9, 49.5]; $z = 10.50$, $p < .001$). We inspected the interaction using follow-up Tukey contrasts correcting for multiple comparisons. Although adults gave higher ratings to Likely Answers (M = 85.81, SD = 18.37) than Unlikely Answers (M = 43.94, SD = 26.02; $p < .001$), they did not differentiate between Non-Answers that were Likely (M = 10.7, SD = 18.4) or Unlikely (M = 10.5, SD = 21.8; $p > .99$). Critically, adults endorsed Unlikely Answers more often than Unlikely Non-Answers ($p < .001$) and also Likely Non-Answers ($p < .001$). See Figure 5a for adults' average ratings.

The results for children were similar (see Figure 6). As in the adult sample, there was a significant effect of Answer Type ($\beta = 1.38$, OR = 3.96, 95% CI [1.81, 8.65], $z = 3.45$, $p < .001$), no main effect of Likelihood ($z = 0.22$, $p = .82$), and a significant Answer Type by Likelihood interaction ($\beta = 3.68$, OR = 39.53, 95% CI [12.28, 127.21]; $z = 6.17$, $p < .001$). Follow-up Tukey contrasts using estimated marginal means indicated that children were more likely to endorse Likely Answers (M = 90%, SE = 4%) than Unlikely Answers (M = 17%, SE = 5%; $p < .001$), but did not differentiate between the Likely Non-Answers (M = 5%, SE = 2%) and Unlikely Non-Answers (M = 5%, SE = 2%; $p > .99$). Children also endorsed Unlikely Answers more often than both Likely Non-Answers ($p = .049$) and Unlikely Non-Answers ($p = .040$).

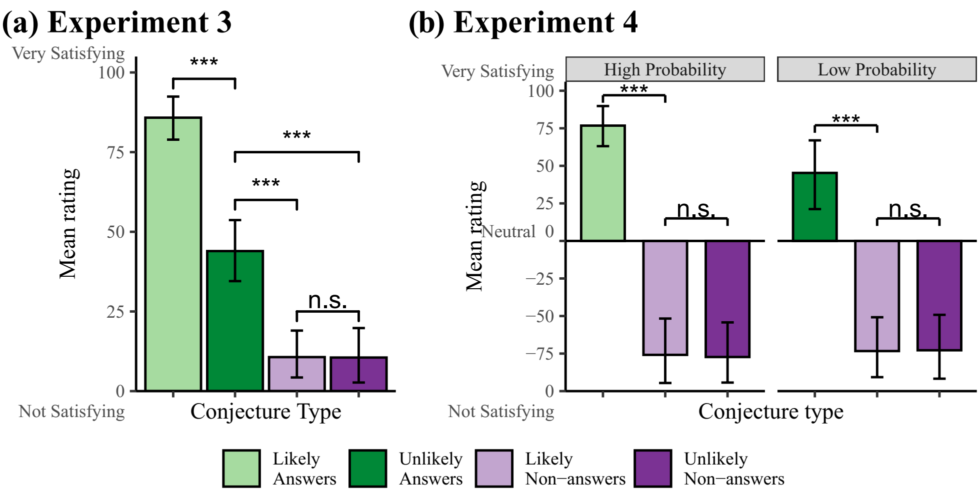

FIGURE 5 Adults' ratings in (a) Experiment 3 (N = 24) and (b) Experiment 4 (N = 24) averaged across the six test trials. When given all responses and a positive rating scale (0–100), adults preferred likely answers to unlikely answers and both of these to non-answers (Experiment 3). The results replicated when given non-answers and *only either* likely *or* unlikely answers, and a scale allowing answers to rate answers either negatively or positively (−100 to 100, Experiment 4). Participants rated unlikely answers nearly as positively as likely ones. Error bars show bootstrapped 95% confidence intervals. (Tukey's HSD, ***p < .001)

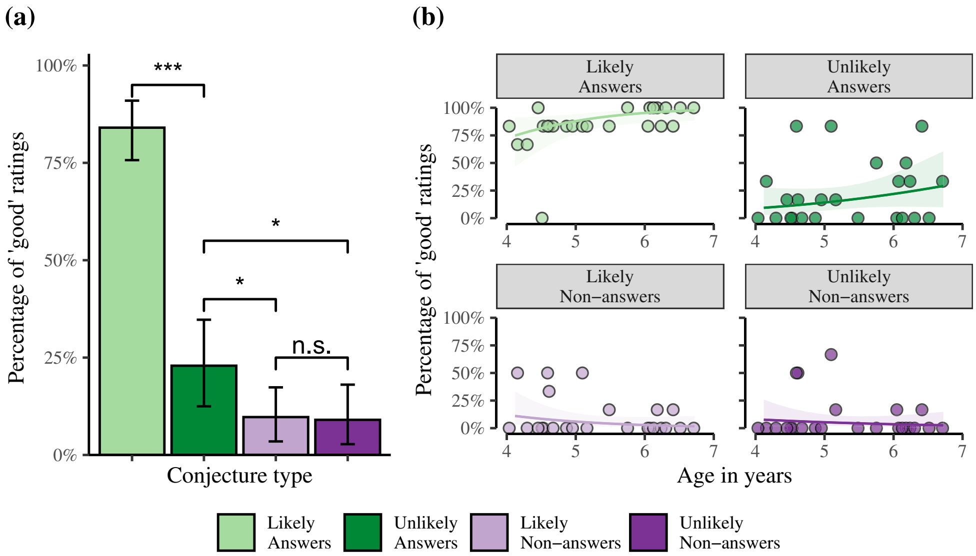

FIGURE 6 Percentage of conjectures rated as good by children in Experiment 3 (N = 24; mean: 5.95, range 4.37–6.71), averaged across the six test trials. Children rated all four conjecture types on each trial. (a) Responses averaged across all participants; (b) The same data by answer type and child's age; each circle represents responses from one child on that answer type. Like adults, children preferred likely answers to unlikely answers and both of these to non-answers. With increasing age, children were more likely to accept Answers and more likely to reject Non-Answers, for both likely and unlikely conjectures. The distinction between Likely and Unlikely Answers decreased with age. Lines show predictions from regression model; shaded regions and error bars show 95% confidence intervals. (Tukey's HSD, *p < .05, ***p < .001)

We also looked at the effect of age on children's responses by adding a covariate of age (in years, mean-centered) to the logistic mixed effects model predicting children's responses and including an age by conjecture interaction. Because we did not have specific hypotheses about how age might interact with children's ratings for each conjecture type, we used a categorical variable of Conjecture Type (with Likely Answers as the reference level) so that estimated coefficients could be directly interpreted as the effect of age on the log odds of endorsing each conjecture type. The model thus included fixed effects of conjecture type, age, and a conjecture type by age interaction, with random intercepts for subject and story. This expanded model explained significant additional variance than the original analysis without age (χ^2 (4) = 14.36, p = .006 by likelihood ratio test), although the overall effect of Conjecture type remained significant (χ^2 (3) = 124, p < .001). Critically, this is qualified by an age by conjecture type interaction (χ^2 (3) = 13.53, p = .004). Inspection of estimated marginal slopes indicated that with increasing age, children were more willing to endorse conjectures that answered the question, regardless of how otherwise plausible they were (Likely Answers: β = 1.02, 95% CI [0.08, 1.96]; Unlikely Answers: β = 0.53, 95% CI [−0.30, 1.37]) and less likely to endorse conjectures that did not answer the question, again regardless of how otherwise plausible they were (Likely Non-Answers: β = −0.77, 95% CI [−1.86, 0.32]; Unlikely Answers: β = −0.43, 95% CI [−1.49, 0.62]). Figure 6b illustrates these marginal effects of age on the probability of endorsing each conjecture type.

Discussion

The results of Experiment 3 suggest that both adults and 4- to 6-year-old children will endorse even otherwise unlikely conjectures as long as they offer potential answers to questions. This is not to say that people are indifferent to the plausibility of conjectures: Both adults and children preferred conjectures involving likely events to those involving unlikely ones. Importantly, however, and consistent with the idea of cognitive pragmatism, participants' evaluations privileged the degree to which a conjecture might answer the question, and only secondarily considered how likely the conjecture might be.

However, although adults rated the unlikely answers as more satisfying than the non-answers, the average rating of unlikely answers was close to the middle of the scale, suggesting that adults might not so much have endorsed as merely been indifferent to the unlikely conjecture. In Experiment 4, we replicate the design of Experiment 3 but use a −100 to +100 rating scale to allow us to distinguish adults' active endorsement of unlikely conjectures from more neutral or negative responses. We also gave both adults and children just one kind of answer, plausible or implausible, to see how participants might evaluate prima facie implausible conjectures when they are not explicitly contrasted with more plausible ones.

EXPERIMENT 4

In Experiment 4, we used a between-participant design for adults and a within-participant design for children. We asked both adults and children to evaluate three conjectures for each question: one low probability conjecture that answered the question (Unlikely Answer) and two conjectures that did not answer the question (i.e., Likely and Unlikely Non-Answers). In adults, we also ran a condition in which participants rated a high probability conjecture that answered the question (Likely Answer) and both Non-Answers. For adults, our primary question of interest was whether there would be any difference in their ratings for Likely versus Unlikely Answers between conditions. For children, our primary aim was to follow up on the age effect and see whether older children would be more likely than younger children to endorse low-probability conjectures that answered the questions at hand (Unlikely Answers). Based on the large effect size in Experiment 3 (Cohen's d = 1.73), we aimed to recruit the same number of participants per sample for this follow-up. Specifically, 12 adults per condition would provide 98% power to detect the same effect size and 80% power to detect a smaller but still substantial effect of d = 1.2. For children we aimed to recruit 12 four- and five-year-olds and 12 six- and seven-year-olds to yield 80% power to detect an age by conjecture type interaction with a large effect size (from Experiment 3: partial η^2 = .25; Cohen's f = .58).

Methods

Participants

Twenty-four adults were recruited and tested via Amazon Mechanical Turk and paid $1.00 for participating. An additional 13 adults were excluded for failing to distinguish good and bad responses on the inclusion items by at least a 10-point spread (n = 12) or failing attention checks (n = 1).

Twenty-four 4- to 8-year-olds (M = 5.32 years, SD = 0.86, range = 4.04–6.71) were recruited and tested as in the preceding experiments. Six additional children participated but were excluded for responding inaccurately on the inclusion questions (N = 5) or failing to complete the study (N = 1).

Materials and procedures

The materials and procedure were the same as in Experiment 3 with three modifications. First, in Experiment 4, participants rated three conjectures per test trial instead of four. Thus, for children, only three puppets were used per trial, for a total of 18 puppets. Second, we modified the rating scale for adult participants to range from −100 ("Extremely Dissatisfying") to +100 ("Extremely Satisfying), with 0 explicitly marked as "Neutral." Children used the same binary response measure from Experiment 2–3, rating each puppet's proposal as a good or not so good answer. Third, we modified the sentence structure of the conjectures. In Experiment 3, we had used one long sentence of the form "Not because [likely answer], but because [conjecture]". In Experiment 4, we separated the denial and the conjecture into two sentences (e.g., "Sally was not excited because it was a birthday invitation. Sally was excited because it was a letter from school saying she had won a story competition.") both to remind participants of the target of the explanation and to make the denial of the likely explanation even more explicit.

Results and discussion

We had two primary questions. First, we wanted to know whether adults would distinguish likely and unlikely conjectures in their evaluations if each conjecture was presented independently. Adults gave higher ratings to Likely Answers (M = 76.8, SD = 15.4) than Unlikely Answers (M = 45.3, SD = 43.6; $t(17.7) = 2.16$, $p = .04$). Importantly, although participants recognized that one answer was more plausible than the other, they treated unlikely conjectures as acceptable answers when the improbable conjecture was presented as the only answer to a question: of the 12 adults rating Unlikely Answers, 10 gave positive ratings, and 3 gave ratings above 90 (comparable to the 5 of 12 adults who gave ratings above 90 in the Likely Answer condition). Consistent with the idea of cognitive pragmatism, adults seemed to value the degree to which a proposal could answer a question above the prior probability of the proposal. See Figure 5b for adults' average responses.

Like adults, children in Experiment 4 endorsed Unlikely Answers (M = 44%) more often than both Likely Non-Answers (M = 15%, Wilcoxon signed rank test $p < .001$) and Unlikely Non-Answers (M = 13%, $p < .001$). Again, however, this was qualified by an age by conjecture type interaction. We built a logistic regression model with the fixed effects of conjecture type, age, and an age by conjecture type interaction. As planned, we analyzed age as a binary variable (ages 4–5 vs. 6–7), although similar results obtain with age coded continuously (see Supporting Information for details). This model explained more variance than either the conjecture only model ($\chi^2(3) = 21.4$, $p < .001$) or the model with conjecture and age but no interaction ($\chi^2(2) = 21.3$, $p < .001$). Inspection of estimated marginal slopes indicated that older children were more likely than younger children to endorse Unlikely Answers ($\beta = 0.34$, $p = .02$), but equally likely to reject conjectures that did not answer the question (Likely Non-Answers: $\beta = -0.091$, $p = .18$; Unlikely Non-Answers: $\beta = 0.088$, $p = .16$). Thus, and in contrast to the idea that younger children might be if anything more drawn to novel and speculative responses than older children, 4- and 5-year-olds were in fact more likely to endorse answers that were probable in themselves whereas the 6- and 7-year-olds were able to represent and evaluate proposals based only on the abstract fit between the question and the answer. Figure 7 shows the distribution of responses by age group.

Finally, we note that in both Experiments 3 and 4, 4- and 5-year-olds correctly endorsed likely answers, correctly rejected both likely and unlikely non-answers, and (unlike adults and older children) sometimes also rejected unlikely answers. By contrast, in Experiment 2, 4- and 5-year-olds tended to rate all responses (facts and conjectures) positively, although they correctly endorsed facts more often for questions that could be answered by those facts in the story and conjectures when they could not. Overall, however, this pattern of results raises the question of why 4- and 5-year-olds did sometimes reject responses in Experiments 3 and 4 but generally (if differentially) endorsed all responses in Experiment 2. We believe two factors might have contributed. First, Experiment 2 never involved any conflict with the children's prior knowledge; this might have made it easier for children to endorse responses across the board. Second, all the responses except the likely response were prefaced with an explicit rejection of at the likely response ("It wasn't a birthday invitation. Sally was excited because

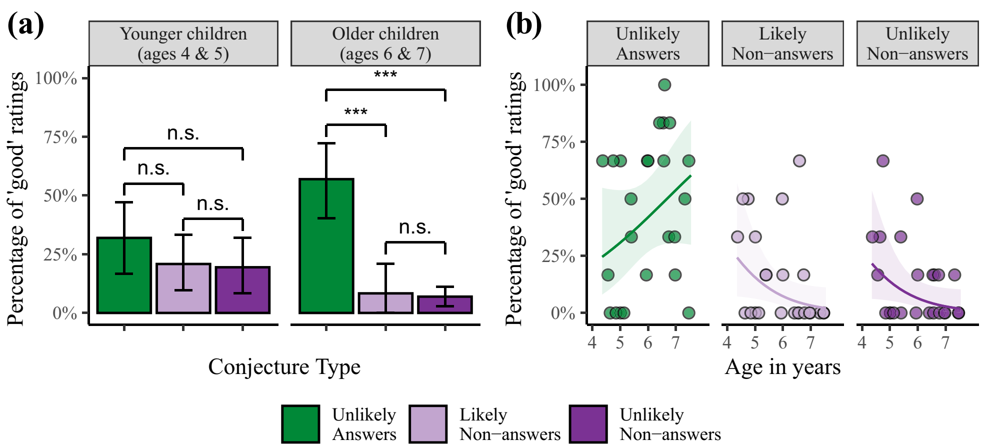

FIGURE 7 Percentage of conjectures rated as good by children in Experiment 4 ($N = 24$; mean: 5.32; range 4.04–6.71), averaged across the six test trials. Children rated all three conjecture types on each trial. (a) Responses averaged across participants by age group. (b) Each circle represents responses from one child on that conjecture type. With increasing age, children were more likely to accept Unlikely Answers and less likely to accept both Likely and Unlikely Non-Answers. Lines show predictions from regression model; shaded regions and error bars show 95% confidence intervals. (Wilcoxon signed rank test, ***$p < .001$)

…"). Since the experimenter effectively modeled rejecting a candidate response, the children might have felt more licensed to do so as well.

GENERAL DISCUSSION

Across four experiments, we found that both adults and children were willing to entertain novel and unverified claims when (and only when) they answered the question at hand. Four- and 5-year-olds endorsed such conjectures in both forced choice (Experiment 1) and independent judgment paradigms (Experiment 2), even when the conjectures were accompanied by explicit expressions of uncertainty (Experiment 2). Adults and 6- and 7-year-old children, but not younger children, further endorsed conjectures that were improbable but provided potential answers to the questions at hand (Experiments 3 and 4). When confronting otherwise unanswered questions, adults and young children judged speculative conjectures not on the basis of evidence for their truth, but instead, on their potential for addressing those unanswered questions *if* they were true.

Our results add an important perspective to the growing literature on trust and testimony, which has largely found that children rationally integrate multiple cues to speaker and information reliability when deciding whether to accept or reject novel claims (for review, see Harris et al., 2018; Koenig et al., 2019; Stephens et al., 2015). In contrast, here we find that participants endorsed novel conjectures despite conflicts with cues to reliability, including verifiability, speaker confidence, and consistency with prior expectations. Our results suggest that the primary driver of learners' judgments was whether the proposed conjecture answered the question at hand. Indeed, in the absence of other potential answers to a question, adults judged implausible conjectures just as favorably as plausible conjectures.

We also found evidence for a developmental trajectory in which older children and adults were more willing to accept low probability conjectures than younger children (Experiments 3–4). Importantly, this was the case only when the low probability conjectures answered the question; participants tended to reject all non-answers. There are at least two mutually compatible explanations for this finding. First, there may be an age-related increase in participants' tolerance for uncertainty and the value they give to getting a question answered—in other words, a move from an empirical stance to a pragmatic stance. This hypothesis is supported by young children's success at recognizing conjectures that answer a question in Experiments 1 and 2. In fact, this developmental finding is compatible with other reports of an age-related increase in the subtlety and sophistication with which children evaluate claims and testimony (for review, see Mills, 2013). Whereas 3-year-olds reject sources with any evidence of historical inaccuracy, 4-year-olds differentiate between sources with 75% versus 25% inaccuracy and preferentially trust the more reliable (Pasquini et al., 2007). Thus, with age, children may learn to balance multiple criteria for deciding which claims to endorse or reject. Future work may explore how children integrate various criteria—empirical, social, and structural—when evaluating novel claims and how properties of the specific question under discussion influences which criteria takes priority.

Second, there may be an age-related improvement in participants' ability to recognize claims that answer a question. Precisely how people assess the satisfactoriness of a conjecture in the absence of any evidence remains an open question; we speculate that people might be broadly sensitive to information contained in the question itself. For instance, we know that by age three, children recognize that interrogative words specify desires for different categories of information (Ervin-Tripp, 1970), such that "who" requires an agent and "where" requires a place; Piaget (1926) further divided "why" questions into subtypes: causal explanation, human motivation, justification, and logical explanations. More generally, beyond linguistic information, the ability to evaluate novel conjectures might be related to developing world knowledge (e.g., in Experiment 3, an understanding of what events might make someone excited or sad) or problem-solving abilities (e.g., in Experiment 1, knowing that you could retrieve out-of-reach objects by knocking them off the platform with a ball). Over time, accumulated linguistic experience, world knowledge, and planning abilities might help children fine-tune their understanding of what different questions require of their answers.

We have contrasted the current work with past work on trust in testimony. Here we show that so long as a response provides a potential answer to a question, children are willing to override known fact (even when communicated confidently) in favor of speculative conjectures (even when advanced uncertainly). However, we do not mean to suggest any fundamental incompatibility between our results and the literature on trust in testimony. Children might well be epistemically vigilant in tracking informants who do and do not provide satisfactory answers to questions, even when all the answers are conjectural. In future research, it might be interesting to see if children track agents' history of answering questions with appropriate (vs. irrelevant) conjectures and whether children use this information to make decisions about whom to address when posing questions likely to have no known answer.

It should be noted that although across these studies we tested a variety of scenarios (made-up monsters and real-life events that differed in plausibility) and used a range of response measures (forced choice and independent rating), we lack direct evidence that children in Experiments 3 and 4 explicitly understood the speculative nature of conjectures. Given that young children can be both overly credulous and overly skeptical of unfamiliar events (Woolley & Ghossainy, 2013), more research is needed to examine children's credulity toward speculative claims. For example, in Experiment 3, we might ask children to explain their ratings and explicitly judge the probability of the conjectures ("Do you think that really happened?") or the speaker's belief in the conjecture ("Does the speaker believe that it really happened?").

We began this line of work by observing that remarkable discoveries often emerge from wild speculations in response to novel, unanswered questions. In science, religion, and everyday life, we give license to conjectures, speculations, and other unverified assumptions when they can play powerful explanatory roles. An important next step is to go beyond looking at how people evaluate claims to looking at how they generate them. Any number of factors might affect people's willingness to engage in speculative reasoning. For example, people might be more willing to generate conjectures in response to scientific questions than religious ones if they believe that accepting mysteries in religion is a sacred value; on the other hand, they might be more willing to generate conjectures in religious contexts than scientific ones if they believe that a variety of answers may be more acceptable in religion than in science (see Liquin et al., 2020). A related question is whether the propensity for speculation and the acceptability of entertaining conjectures differs by domain or situational and social demands. For example, people might differ in their willingness or ability to generate useful speculations, perhaps modulated by their own prior knowledge and an estimation of their own and others' expertise.

While little is known about the origins of speculation in early childhood, it is intuitively clear that children begin attempting to answer questions as soon as they can ask them. Children's conjectures are necessarily limited by their world knowledge and are often wrong with respect to the facts. Nonetheless, a remarkable feature of children's speculations is that their answers are "at least wrong." For instance, a 3-year-old with whom the senior author is well acquainted once speculated in response to the announcement that everyone had to turn off their cell phones when the plane took off that this was because "Planes are noisy and you wouldn't be able to hear if you talked on the phone." This answer is wrong but it is "at least wrong." Consider the infinite variety of things she could have said that would have been, even if factually correct, simply irrelevant (e.g., "Planes are silver and phones are silver"; "planes are big and phones are small"; see Schulz, 2012b for discussion) This early emerging ability to map the structural form of an answer onto a question might be critical to how we can generate new hypotheses. By placing a high value on getting our questions answered, and by readily entertaining new ideas before obtaining evidential support, we may be motivated to explore and inquire in ways that can generate unexpected discoveries.

In conclusion, the results point to a willingness to entertain potentially explanatory but unverified speculations beginning in early childhood. Children as young as four knowingly rejected known information in favor of the unknown, but only when the conjectures addressed an otherwise unanswered question. When evidence is unavailable, children and adults willingly drop their empirical stance to consider the potential value of new ideas, even when these proposals are advanced by uncertain speakers and in tension with prior expectations, in order to find possible solutions to otherwise unresolved problems.

ACKNOWLEDGMENTS

We thank Cindy Zhou, Jaemarie Solyst, Rucha Kelkar, Grace Cowles, and Jenny Zhu for their help with data collection and annotation. We are grateful to the Boston Children's Museum and all the families for participating in this research. This research is funded by NSF Center for Brains Minds and Machines (#1231216).

ORCID

Junyi Chu https://orcid.org/0000-0003-2778-6317

REFERENCES

Bascandziev, I., & Harris, P. L. (2014). In beauty we trust: Children prefer information from more attractive informants. *British Journal of Developmental Psychology*, *32*, 94–99. https://doi.org/10.1111/bjdp.12022

Birch, S. A., Severson, R. L., & Baimel, A. (2020). Children's understanding of when a person's confidence and hesitancy is a cue to their credibility. *PLoS One*, *15*, e0227026. https://doi.org/10.1371/journal.pone.0227026

Bonawitz, E. B., & Lombrozo, T. (2012). Occam's rattle: Children's use of simplicity and probability to constrain inference. *Developmental Psychology*, *48*, 1156–1164. https://doi.org/10.1037/a0026471

Bridgers, S., Buchsbaum, D., Seiver, E., Griffiths, T. L., & Gopnik, A. (2016). Children's causal inferences from conflicting testimony and observations. *Developmental Psychology*, *52*, 9–18. https://doi.org/10.1037/a0039830

Brosseau-Liard, P. E., & Birch, S. A. (2010). 'I bet you know more and are nicer too!': What children infer from others' accuracy. *Developmental Science*, *13*, 772–778. https://doi.org/10.1111/j.1467-7687.2009.00932.x

Brosseau-Liard, P. E., Cassels, T. G., & Birch, S. A. (2014). You seem certain but you were wrong before: Developmental change in preschoolers' relative trust in accurate versus confident speakers. *PLoS One*, *9*, e108308. https://doi.org/10.1371/journal.pone.0108308

Butler, L. P., Schmidt, M. F., Tavassolie, N. S., & Gibbs, H. M. (2018). Children's evaluation of verified and unverified claims. *Journal of Experimental Child Psychology*, *176*, 73–83. https://doi.org/10.1016/j.jecp.2018.07.007

Callanan, M. A., & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development*, *7*, 213–233. https://doi.org/10.1016/0885-2014(92)90012-G

Chan, C. C., & Tardif, T. (2013). Knowing better: The role of prior knowledge and culture in trust in testimony. *Developmental Psychology*, *49*, 591. https://doi.org/10.1037/a0031336

Chen, E. E., Corriveau, K. H., & Harris, P. L. (2012). Children trust a consensus composed of outgroup members-but do not retain that trust. *Child Development*, *84*, 269–282. https://doi.org/10.1111/j.1467-8624.2012.01850.x

Chouinard, M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*, 1–112. https://doi.org/10.1111/j.1540-5834.2007.00412.x

Clark, C. D. (1995). *Flights of fancy, leaps of faith: Children's myths in contemporary America*. University of Chicago Press.

Clément, F. (2010). To trust or not to trust? Children's social epistemology. *Review of Philosophy and Psychology*, *1*, 531–549. https://doi.org/10.1007/s13164-010-0022-3

Clément, F., Koenig, M., & Harris, P. (2004). The ontogenesis of trust. *Mind & Language*, *19*, 360–379. https://doi.org/10.1111/j.0268-1064.2004.00263.x

Corriveau, K. H., & Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. *Developmental Psychology*, *46*, 437–445. https://doi.org/10.1037/a0017553

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*, 1827–1835. https://doi.org/10.1111/cdev.12240

Danovitch, J. H., & Keil, F. C. (2004). Should you ask a fisherman or a biologist?: Developmental shifts in ways of clustering knowledge. *Child Development*, *75*, 918–931. https://doi.org/10.1111/j.1467-8624.2004.00714.x

DiYanni, C. J., Corriveau, K. H., Kurkul, K., Nasrini, J., & Nini, D. (2015). The role of consensus and culture in children's imitation of inefficient actions. *Journal of Experimental Child Psychology*, *137*, 99–110. https://doi.org/10.1016/j.jecp.2015.04.004

Einav, S., & Robinson, E. J. (2011). When being right is not enough: Four-year-olds distinguish knowledgeable informants from merely accurate informants. *Psychological Science*, *22*, 1250–1253. https://doi.org/10.1177/0956797611416998

Elashi, F. B., & Mills, C. M. (2014). Do children trust based on group membership or prior accuracy? The role of novel group membership in children's trust decisions. *Journal of Experimental Child Psychology*, *128*, 88–104. https://doi.org/10.1016/j.jecp.2014.07.003

Ervin-Tripp, S. M. (1970). Discourse agreement: How children answer questions. In J. R. Hayes (Ed.), *Cognition and the development of knowledge* (pp. 79–107). Wiley.

Flavell, J. H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 244–267). Cambridge University Press.

Frazier, B. N., Gelman, S. A., & Wellman, H. M. (2009). Preschoolers' search for explanatory information within adult-child conversation. *Child Development*, *80*, 1592–1611. https://doi.org/10.1111/j.1467-8624.2009.01356.x

Fusaro, M., Corriveau, K. H., & Harris, P. L. (2011). The good, the strong, and the accurate: Preschoolers' evaluations of informant attributes. *Journal of Experimental Child Psychology*, *110*, 561–574. https://doi.org/10.1016/j.jecp.2011.06.008

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

Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of omission: Children selectively explore when teachers are underinformative. *Cognition*, *132*, 335–341. https://doi.org/10.1016/j.cognition.2014.04.013

Harris, P. L. (2000). *Understanding children's worlds: The work of the imagination*. Blackwell.

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251–273. https://doi.org/10.1146/annurev-psych-122216-011710

Huh, M., Grossmann, I., & Friedman, O. (2019). Children show reduced trust in confident advisors who are partially informed. *Cognitive Development*, *50*, 49–55. https://doi.org/10.1016/j.cogdev.2019.02.003

Jaswal, V. K. (2004). Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child Development*, *7*, 1871–1885. https://doi.org/10.1111/j.1467-8624.2004.00822.x

Jaswal, V. K., & Malone, L. S. (2007). Turning Believers into Skeptics: 3-year-olds' sensitivity to cues to speaker credibility. *Journal of Cognition and Development*, *8*, 263–283. https://doi.org/10.1080/15248370701446392

Johnston, A. M., Johnson, S. G. B., Koven, M. L., & Keil, F. C. (2016). Little Bayesians or little Einsteins? Probability and explanatory virtue in children's inferences. *Developmental Science*, *20*, 1–14. https://doi.org/10.1111/desc.12483

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, *14*, 106–111. https://doi.org/10.1111/j.1467-7687.2010.00965.x

Koenig, M., Clément, F., & Harris, P. L. (2004). Trust in testimony: children's use of true and false statements. *Psychological Science*, *14*, 694–698. https://doi.org/10.1111/j.0956-7976.2004.00742.x

Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition*, *87*, 179–208. https://doi.org/10.1016/s0010-0277(03)00002-7

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*, 1261–1277. https://doi.org/10.1111/j.1467-8624.2005.00849.x

Koenig, M. A., & Jaswal, V. K. (2011). Characterizing children's expectations about expertise and incompetence: Halo or pitchfork effects? *Child Development*, *82*, 1634–1647. https://doi.org/10.1111/j.1467-8624.2011.01618.x

Koenig, M. A., Tiberius, V., & Hamlin, J. K. (2019). Children's judgments of epistemic and moral agents: From situations to intentions. *Perspectives on Psychological Science*, *14*, 344–360. https://doi.org/10.1177/1745691618805452

Koenig, M. A., & Woodward, A. L. (2010). Sensitivity of 24-month-olds to the prior inaccuracy of the source: Possible mechanisms. *Developmental Psychology*, *46*, 815–826. https://doi.org/10.1037/a0019664

Kominsky, J. F., Langthorne, P., & Keil, F. C. (2016). The better part of not knowing: Virtuous ignorance. *Developmental Psychology*, *52*, 31–45. https://doi.org/10.1037/dev0000065

Kurkul, K. E., & Corriveau, K. H. (2017). Question, explanation, follow-up: A mechanism for learning from others? *Child Development*, *89*, 280–294. https://doi.org/10.1111/cdev.12726

Kushnir, T. (2013). How children learn from and about people: The fundamental link between social cognition and statistical evidence. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 191–196). Oxford University Press.

Kushnir, T., Vredenburgh, C., & Schneider, L. A. (2013). "Who can help me fix this toy?" The distinction between causal knowledge and word knowledge guides preschoolers' selective requests for information. *Developmental Psychology*, *49*, 446–453. https://doi.org/10.1037/a0031649

Landrum, A. R., Mills, C. M., & Johnston, A. M. (2013). When do children trust the expert? Benevolence information influences children's trust more than expertise. *Developmental Science*, *16*, 622–638. https://doi.org/10.1111/desc.12059

Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development*, *81*, 929–944. https://doi.org/10.1111/j.1467-8624.2010.01443.x

Legare, C. H., Mills, C. M., Souza, A. L., Plummer, L. E., & Yasskin, R. (2013). The use of questions as problem-solving strategies during early childhood. *Journal of Experimental Child Psychology*, *114*, 63–76. https://doi.org/10.1016/j.jecp.2012.07.002

Lillard, A. (2001). Pretend play as twin earth: A social-cognitive analysis. *Developmental Review*, *21*, 495–531. https://doi.org/10.1006/drev.2001.0532

Lillard, A., Pinkham, A. M., & Smith, E. (2011). Pretend play and cognitive development. In U. Goswami (Ed.), *Pretend play and cognitive development* (pp. 285–311). Wiley-Blackwell.

Liquin, E., Metz, S. E., & Lombrozo, T. (2020). Science demands explanation, religion tolerates mystery. *Cognition*, *204*, 104398. https://doi.org/10.1016/j.cognition.2020.104398

Lombrozo, T. (2011). The instrumental value of explanations. *Philosophy Compass*, *6*, 539–551. https://doi.org/10.1111/j.1747-9991.2011.00413.x

Ma, L., & Ganea, P. A. (2010). Dealing with conflicting information: Young children's reliance on what they see versus what they are told. *Developmental Science*, *13*, 151–160. https://doi.org/10.1111/j.1467-7687.2009.00878.x

MacDonald, K., Schug, M., Chase, E., & Barth, H. (2013). My people, right or wrong? Minimal group membership disrupts preschoolers' selective trust. *Cognitive Development*, *28*, 247–259. https://doi.org/10.1016/j.cogdev.2012.11.001

Magid, R., Sheskin, M., & Schulz, L. (2015). Imagination and the generation of new ideas. *Cognitive Development*, *34*, 99–110. https://doi.org/10.1016/j.cogdev.2014.12.008

Mercier, H., Bernard, S., & Clément, F. (2014). Early sensitivity to arguments: How preschoolers weight circular arguments. *Journal of Experimental Child Psychology*, *125*, 102–109. https://doi.org/10.1016/j.jecp.2013.11.011

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology*, *49*, 404–418. https://doi.org/10.1037/a0029500

Mills, C. M., Danovitch, J. H., Rowles, S. P., & Campbell, I. L. (2017). Children's success at detecting circular explanations and their interest in future learning. *Psychonomic Bulletin Review*, *24*, 1465–1477. https://doi.org/10.3758/s13423-016-1195-2

Mills, C. M., Legare, C. H., Bills, M., & Mejias, C. (2010). Preschoolers use questions as a tool to acquire knowledge from different sources. *Journal of Cognition and Development*, *11*, 533–560. https://doi.org/10.1080/15248372.2010.516419

Mills, C. M., Legare, C. H., Grant, M. G., & Landrum, A. R. (2011). Determining who to question, what to ask, and how much information to ask for: The development of inquiry in young children. *Journal of Experimental Child Psychology*, *110*, 539–560. https://doi.org/10.1016/j.jecp.2011.06.003

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, *60*, 167–171. https://doi.org/10.2307/1131082

Morgan, T. J. H., Laland, K. N., & Harris, P. L. (2014). The development of adaptive conformity in young children: Effects of uncertainty and consensus. *Developmental Science*, *18*, 511–524. https://doi.org/10.1111/desc.12231

Nguyen, S. P. (2012). The role of external sources of information in children's evaluative food categories. *Infant and Child Development*, *21*, 216–235. https://doi.org/10.1002/icd.745

Nurmsoo, E., & Robinson, E. J. (2009). Children's trust in previously inaccurate informants who were well or poorly informed: When past errors can be excused. *Child Development*, *80*, 23–27. https://doi.org/10.1111/j.1467-8624.2008.01243.x

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*, 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Piaget, J. (1926). *The Language and Thought of the Child*. Harcourt.

Piaget, J. (1929). *The child's conception of the world*. Routledge.

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2015). The effects of collaboration and minimal-group membership on children's prosocial behavior, liking, affiliation, and trust. *Journal of Experimental Child Psychology*, *139*, 161–173. https://doi.org/10.1016/j.jecp.2015.05.008

Popper, K. R. (1959). *The logic of scientific discovery*. Basic Books.

Prentice, N. M., & Gordon, D. A. (1987). Santa Claus and the Tooth Fairy for the Jewish child and parent. *The Journal of Genetic Psychology*, *148*, 139–151. https://doi.org/10.1080/00221325.1987.9914544

Prentice, N. M., Manosevitz, M., & Hubbs, L. (1978). Imaginary figures of early childhood: Santa claus, easter bunny, and the tooth fairy. *American Journal of Orthopsychiatry*, *48*, 618. https://doi.org/10.1111/j.1939-0025.1978.tb02566.x

Reyes-Jaquez, B., & Echols, C. H. (2013). Developmental differences in the relative weighing of informants' social attributes. *Developmental Psychology*, *49*, 602–613. https://doi.org/10.1037/a0031674

Ronfard, S., Zambrana, I. M., Hermansen, T. K., & Kelemen, D. (2018). Question-asking in childhood: A review of the literature and a framework for understanding its development. *Developmental Review*, *49*, 101–120. https://doi.org/10.1016/j.dr.2018.05.002

Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology*, *52*, 2159–2173. https://doi.org/10.1037/dev0000240

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*, 1054–1070. https://doi.org/10.1111/1467-8624.00334

Sabbagh, M. A., & Shafman, D. (2009). How children block learning from ignorant speakers. *Cognition*, *112*, 415–422. https://doi.org/10.1016/j.cognition.2009.06.005

Schulz, L. E. (2012a). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences*, *16*, 382–389. https://doi.org/10.1016/j.tics.2012.06.004

Schulz, L. E. (2012b). Finding new facts; thinking new thoughts. *Advances in Child Development and Behavior*, *43*, 269–294. https://doi.org/10.1016/b978-0-12-397919-3.00010-1

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. https://doi.org/10.1037/0012-1649.43.5.1124

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science*, *22*, 878–880. https://doi.org/10.1177/0956797611411584

Shafto, P., Eaves, B., Navarro, D. J., & Perfors, A. (2012). Epistemic trust: Modeling children's reasoning about others' knowledge and intent. *Developmental Science*, *15*, 436–447. https://doi.org/10.1111/j.1467-7687.2012.01135.x

Sharon, T., & Woolley, J. D. (2004). Do monsters dream? Young children's understanding of the fantasy/reality distinction. *British Journal of Developmental Psychology*, *22*, 293–310. https://doi.org/10.1348/026151004323044627

Sobel, D. M., & Corriveau, K. H. (2010). Children monitor individuals' expertise for word learning. *Child Development*, *81*, 669–679. https://doi.org/10.1111/j.1467-8624.2009.01422.x

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review*, *120*, 779–797. https://doi.org/10.1037/a0034191

Sobel, D. M., & Macris, D. M. (2013). Children's understanding of speaker reliability between lexical and syntactic knowledge. *Developmental Psychology*, *49*, 523–532. https://doi.org/10.1037/a0029658

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*, 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Stephens, E., Suarez, S., & Koenig, M. (2015). Early testimonial learning. In J. B. Benson (Ed.), *Advances in child development and behavior* (pp. 151–183). Elsevier.

Taylor, M., Cartwright, B. S., & Carlson, S. M. (1993). A developmental investigation of children's imaginary companions.

Developmental Psychology, *29*, 276. https://doi.org/10.1037/0012-1649.29.2.276

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

Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology*, *47*, 1065–1077. https://doi.org/10.1037/a0023273

Tizard, B., & Hughes, M. (1984). *Children learning at home and in school*. Fontana.

Tsividis, P., Tenenbaum, J. B., & Schulz, L. E. (2015). Hypothesis-Space constraints in causal learning. In *37th Annual Proceedings of the Cognitive Science Society*, Pasadena, CA, pp. 2434–2439.

VanderBorght, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development: An International Journal of Research and Practice*, *18*, 61–71. https://doi.org/10.1002/icd.591

Walker, C. M., Bonawitz, E. B., & Lombrozo, T. (2017). Effects of explaining on children's preference for simpler hypotheses. *Psychonomic Bulletin Review*, *24*, 1538–1547. https://doi.org/10.3758/s13423-016-1144-0

Walker, C. M., Gopnik, A., & Ganea, P. A. (2015). Learning to learn from stories: Children's developing sensitivity to the causal structure of fictional worlds. *Child Development*, *86*, 310–318. https://doi.org/10.1111/cdev.12287

Weisberg, D. S., & Gopnik, A. (2013). Pretense, counterfactuals, and Bayesian causal models: Why what is not real really matters. *Cognitive Science*, *37*, 1368–1381. https://doi.org/10.1111/cogs.12069

Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, *33*, 341–356. https://doi.org/10.1016/j.dr.2013.08.002

Woolley, J. D. (1995). The fictional mind: Young children's understanding of imagination, pretense, and dreams. *Developmental Review*, *15*, 172. https://doi.org/10.1006/drev.1995.1008

Woolley, J. D., & Ghossainy, E. (2013). Revisiting the fantasy-reality distinction: Children as naïve skeptics. *Child Development*, *84*, 1496–1510. https://doi.org/10.1111/cdev.12081

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*, 245–272. https://doi.org/10.1037/0033-295X.114.2.245

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Chu, J., & Schulz, L. E. (2021). Children selectively endorse speculative conjectures. *Child Development*, 00, 1–19. https://doi.org/10.1111/cdev.13647