Cognition 276 (2026) 106668



Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit



Children expect prior belief strength to constrain belief revision

Pearl Han Li [a,*,1], Laura K. Soter [b,*,1], Tamar Kushnir [c]

[a] Department of Psychology, University of Wisconsin-Madison, 1202 W Johnson Street, Madison, WI 53706, United States of America
[b] Department of Philosophy, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada
[c] Department of Psychology, Duke University, 417 Chapel Drive, Durham, NC 27708, United States of America

ARTICLE INFO

Keywords:
Belief revision
Theory of mind
Prior strength
Belief updating

ABSTRACT

Across two preregistered studies ($N = 199$), we investigate whether children's developing theory of mind incorporates assumptions that people revise their beliefs rationally. Specifically, we focus on the role of *prior belief strength* (magnitude of and degree of certainty in the belief) in shaping children's judgments and predictions about others' belief revision. We presented four- to eight-year-old children with pairs of characters who hold the same belief, but differ in the strength and certainty of their belief. In Study 1, we found that children judge it possible to change both strong and weakly held beliefs—but by age five, children judge strongly held beliefs to be both less possible to change, and more difficult to change than weakly held ones. Additionally, children's open-ended explanations predominantly cited evidence as the reason for belief change. In Study 2, we tested whether children use prior belief strength to predict responses to counterevidence. After both characters encountered identical disconfirming testimony from a group of peers, children by age five predicted that characters with weak priors would revise their beliefs more readily than those with strong priors. In both studies, we found these effects of prior belief across diverse belief domains. Together, these studies offer evidence that children's developing theory of mind–in particular, their understanding of belief change–incorporates an assumption that beliefs are graded states whose revision is constrained by prior strength.

Imagine predicting whether two different people would change their minds when presented with counterevidence. In the first: a friend mentions they think the zoo has pandas, but they are not very sure. You visited yesterday and the zookeeper told you there are no pandas. You share this information, fully expecting your friend to revise their belief: their initial belief was weak, and your evidence is strong. In the second: a colleague who visits the zoo regularly is confident the zoo has pandas. You vaguely recall someone mentioning there were no pandas, but you are not certain. Here, you would not expect your colleague to change their mind: their belief is well-founded, and your evidence is weak. These intuitions reflect a fundamental principle articulated by epistemological theories of rational belief updating: whether and how people revise their beliefs ought to depend on how strongly they hold their original belief and the strength of any new evidence they encounter (Gärdenfors & Makinson, 1988; Lin, 2024; Titelbaum, 2022).

Separately from philosophical claims about how beliefs ought to change is the question of how we (as ordinary reasoners) believe beliefs do change. Our folk intuitions about dynamics of belief revision can be thought of as a component of our developing theory of mind (Wellman, 2014, 2018). Understanding children's emerging understanding of belief centrally involves their understanding about how beliefs change. Yet no prior work has directly tested whether children expect prior belief strength to constrain others' belief revision.

This question lies at the heart of how children think about and interact with other minds. The answer bears on how children conceive of belief itself: whether they represent beliefs as graded states that can be held more or less strongly, and how they reason about the conditions under which beliefs will change. But the stakes extend well beyond children's representation of belief as a mental state. If children hold expectations about rational belief revision, these expectations also scaffold a broad range of children's social behavior and evaluations of others. For instance, such expectations can guide whether children try to change someone's mind and how much evidence they think they need to offer; when they decide it is reasonable to stop trying; and when they judge someone as open- or close-minded, gullible, or rational.

Despite these important implications, no prior work has directly tested whether children understand that others' belief revision is shaped by the strength of their initial belief. In this paper, we present two

* Corresponding authors.
E-mail addresses: pearlhan.li@wisc.edu (P.H. Li), lksoter@yorku.ca (L.K. Soter).
[1] Indicates joint first-authorship.

https://doi.org/10.1016/j.cognition.2026.106668
Received 7 November 2025; Received in revised form 13 July 2026; Accepted 16 July 2026
Available online 19 July 2026
0010-0277/© 2026 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

studies that address this question.

1.1. Rational belief updating

Belief updating refers to how an agent changes their beliefs when they acquire new relevant information. Standard theories of rational belief revision (Gärdenfors & Makinson, 1988; Lin, 2024; Titelbaum, 2022) identify two key components that determine how an agent ought to update beliefs: (1) the strength of the prior belief (how much confidence or credence one has in the target claim at the start) and (2) the strength of the new evidence (how much some new piece of information weighs in favor of or against the target claim). This relationship can be formally modeled in a Bayesian framework, but can also be understood more intuitively: if someone believes something strongly (has a high prior) and encounters weak counterevidence to their belief, they should not reduce their initial belief substantially; in contrast, if they believe something only weakly and encounter strong counterevidence, they should revise their belief more dramatically. This framework suggests that strong prior beliefs function as *epistemic constraints* on belief revision—they create resistance to change that can only be overcome by sufficiently compelling evidence. Importantly, this framework is domain general, capturing the dynamics of rational belief revision for any kind of belief content. While the specific criteria for confidence or evidential strength may vary across domains, the basic conceptual framework remains constant.

Most psychological research adopting this rational model has explored whether adults and children actually revise their beliefs in the way it prescribes. Such work has found evidence that people's belief revision practices often do cohere with the rational model: that is, both adults and children update their beliefs in a way that is proportional to their priors and the strength of the novel evidence they encounter (Jern et al., 2014; Liu & Xu, 2022, 2023; Masnick et al., 2017; Tenenbaum et al., 2006; Xu & Tenenbaum, 2007). Many developmental theorists argue that children's learning about the world proceeds in a broadly rational and evidence-responsive manner (Gopnik & Wellman, 2012; Tenenbaum et al., 2011; Xu & Kushnir, 2013). Children's belief updating is sensitive to the probabilistic strength of novel evidence (Kimura & Gopnik, 2019; Kushnir & Gopnik, 2005, 2007; Macris & Sobel, 2017; Schleihauf et al., 2022), and much work has shown that how such counterevidence is integrated depends on the strength of their initial beliefs. For example, in the domain of causal learning, children's integration of new causal information is sensitive to prior base rate beliefs and existing domain knowledge (Griffiths et al., 2011; Kushnir & Gopnik, 2007; Schulz & Gopnik, 2004; Sobel et al., 2004). More broadly, children develop intuitive theories that they revise in response to evidence, with the degree of revision depending on their prior confidence (Gopnik & Wellman, 2012). Similarly, in social learning, children modulate their acceptance of testimony based on a speaker's knowledge and reliability (Jaswal & Neely, 2006; Koenig & Harris, 2005a, 2005b; Kushnir et al., 2008; Sabbagh & Baldwin, 2001), suggesting sensitivity to the evidential quality of the information they receive (Sobel & Kushnir, 2013).

1.2. From rational revision to theory of mind

This research compellingly shows that children's learning is often epistemically rational, in that belief updating is proportional to prior strength and evidence strength. But it does not yet tell us when and whether children come to recognize and/or expect such rational belief updating in others; after all, the fact that (even very young) children show the capacity to behave in a particular way is no guarantee that they have developed an expectation that other people will or should do the same. Does children's theory of mind include a representation of belief revision as rational (i.e. as determined by the degree of prior belief strength and evidence strength)?

Decades of theory of mind research (Goodman et al., 2006; Rakoczy, 2022; Wellman, 2014, 2018) have provided crucial foundations for answering this question. In the domain of action, infants and preschool-age children represent others as rational agents whose behavior is produced by their beliefs, goals and desires, and physical/environmental constraints in predictable ways (Gergely & Csibra, 2003; Jara-Ettinger et al., 2015; Wellman, 2014; Luo & Baillargeon, 2007; Liu et al., 2017; Sommerville & Woodward, 2005; Eason et al., 2018). Further, infants and preschoolers make inferences about agents' goals and preferences based on rational statistical constraints on behavior (Gweon et al., 2010; Kushnir et al., 2010; Wellman et al., 2016). On the basis of such findings (and others), some have argued that infants and young children rely on a "naïve utility calculus" through which they represent others as rational agents who maximize utility given preferences, motivations, and costs (Jara-Ettinger et al., 2015; Jara-Ettinger et al., 2017).

This body of work establishes that children's theory of mind incorporates an assumption that agents are rational in the domain of action. But does such an assumption also extend to the domain of belief? If children's theory of mind casts others (at least by default) as *epistemically rational believers* in the same way that their intuitive theory of agency casts others (at least by default) as rational actors, they should similarly be able to offer predictions, judgments, and explanations about others' belief revision (Gerstenberg & Tenenbaum, 2017; Mahr & Csibra, 2021).

There is much we already know about children's understanding of belief and belief change. First and most basically, children understand that beliefs can change. Even three-year-olds talk about learning (Sobel et al., 2007) and understand that people can come to have knowledge in the future that they currently lack (Atance & Caza, 2018). Children also become sensitive to the possibility of representational change around age four (Gopnik & Astington, 1988). So although children's full understanding of learning (especially their metacognitive understanding of their *own* learning) continues to develop into middle childhood (Sobel et al., 2007; Sobel & Letourneau, 2015; Wang, 2015), this basic understanding that beliefs can change and knowledge can be acquired emerges from a very young age. Second, young children understand that beliefs can vary in strength at least in a rudimentary form: that is, they distinguish between expressions of confidence versus uncertainty, and they can use this information in deciding who to learn from (Baer et al., 2025; Kushnir & Koenig, 2017; Sabbagh & Baldwin, 2001). By age five, for example, children use cues about an informants' certainty or uncertainty to shape the inferences they make in the domain of causal learning (McLoughlin et al., 2021). Third, young children understand that beliefs are shaped by evidence. Children as young as three recognize that epistemic or evidential reasons are better reasons for belief than practical reasons (Pesch et al., 2022), and that people cannot simply adopt a belief that directly contradicts their evidence (Confer et al., 2024). Fourth, children also understand that evidence can vary in strength, and they can use this to shape both their own belief-updating and their predictions of others' belief-updating (Langenhoff et al., 2023; Magid et al., 2018; Young et al., 2012).

Together, these findings show that children are sensitive to the key components that would be needed to support an assumption of rational belief revision: that beliefs can change, that they change in response to evidence, and that beliefs can vary in strength. But no prior research has put all of these pieces together to examine whether children's theory of mind includes expectations about the rationality of belief revision dynamics. In particular, no study has investigated whether children understand that prior belief strength interacts with evidence to shape how beliefs will update. This is the question we test in the present work.

1.3. Prior strength as a constraint on belief updating

Our central question, and the one that has not been directly investigated, is whether children understand that prior strength shapes belief updating–i.e., that it takes more evidence to update a strongly held belief than a weakly held one. Prior work has shown that young children understand not only physical constraints, but also *mental* constraints on

what can be done (Browne & Woolley, 2004; Kushnir et al., 2015; Sobel, 2004). This includes an appreciation for *epistemic* constraints; for instance, four-year-olds understand that someone is not free to "act otherwise" if their actions are determined by a lack of requisite knowledge (Chernyak et al., 2013; Kushnir et al., 2015). In light of this, we can frame our central research question as follows: do children conceive of prior belief strength as an epistemic constraint that systematically influences the attributed possibility, difficulty, and likelihood of belief change?

Two recent lines of work are particularly relevant to this question. First, Confer et al. (2024) show that young children understand epistemic and evidential constraints on belief. Children evaluated characters who held beliefs that were either clearly contradicted by evidence (e.g., believing it was raining while seeing sunshine) or consistent with it. Five-year-olds recognized that people can revise their beliefs upon encountering strong counterevidence for their initial beliefs, but that they cannot simply decide to change their beliefs when strong evidence supports it. These findings suggest that children construe beliefs as evidence-constrained mental states. Second, Magid et al. (2018) examined whether children use *evidence quality* to predict others' belief revision. After learning about a puppet with initial beliefs about the location of preferred toys, four- to six-year-old children expected the puppet to update his beliefs when he received strong, randomly sampled evidence, but not when evidence was weak or selectively sampled. This work demonstrates that children are sensitive to evidence strength when predicting belief revision.

These findings establish that children treat evidence strength as a constraint on belief revision. To add to this body of work, we directly investigate the role of *prior belief strength* in shaping children's judgments and expectations about belief change. The work reviewed thus far suggests that children understand that beliefs can vary in strength (at least, along a simple dimension of certainty vs uncertainty, e.g. Sabbagh & Baldwin, 2001), and that beliefs should respond to and be constrained by evidence (Confer et al., 2024; Magid et al., 2018; Pesch et al., 2022). What remains untested is whether children integrate these components and have expectations that beliefs are graded states held with varying degrees of certainty, whose revision is constrained by prior strength.

1.4. The present studies

The present studies investigate whether children's developing theory of mind incorporates assumptions that people revise their beliefs rationally given prior beliefs. We test whether prior strength influences both the perceived possibility and difficulty of belief change (Study 1), as well as the likelihood that someone will actually revise their belief in response to counterevidence (Study 2). Across studies, we manipulate prior strength by combining two cues: direct statements of the magnitude of belief strength (characters believe "a lot" or "a little") and also a more indirect cue indicating the characters' certainty/uncertainty ("very sure" "not so sure", respectively). Because prior belief strength is the central manipulation in both studies, we used both cues to ensure that children clearly understood the distinction between strongly and weakly held prior beliefs.

Study 1 examines two facets of children's reasoning about belief change. First, we test whether children think prior strength shapes the *possibility* of belief change. Research shows that young children conceptualize mental states as powerful constraints on what is possible: for instance, they do not recognize that strong desires (Kushnir et al., 2015) or norms (Chernyak et al., 2019, 2013; Shtulman & Phillips, 2018) can be overridden until around age six. If children similarly conceptualize strong prior beliefs as constraints on what is possible, younger children may judge it to be less possible—or even impossible—to change a strongly held belief. We predict that if children's emerging understanding of belief includes sensitivity to the epistemic constraint of prior strength, they will judge weakly held beliefs as more possible to change than strongly held ones.

Second, we test whether children think prior strength shapes the *difficulty* of belief change. Even if children judge both types of beliefs as possible to change, changing a strongly held belief should be seen as harder. As children develop, they acquire increasing appreciation for the difficulty of some mental activities: older children (eight- and thirteen-year-olds) understand better than younger children (five-year-olds) that thoughts are not always fully controllable (Flavell et al., 1998), and older children (10-year-olds) can more reliably distinguish between mental states that are hard versus easy to control than younger children (five-year-olds; Flavell & Green, 1999). We predict that if children understand how prior strength shapes belief revision, they should judge strongly held beliefs as more difficult to change than weakly held ones.

Study 1 also explores children's *explanations* for belief revision. In the domain of action, important evidence for children's intuitive theories of rational agency comes from their explanations: when asked why someone acted as they did, children appeal to the actor's beliefs, desires, and constraints (Bartsch & Wellman, 1989), demonstrating an understanding of the causal structure of action. Here, we investigate whether children's *explanations* of others' belief-revision similarly reveal their understanding of the mechanisms of belief change. In particular, if children's understanding of beliefs includes the idea that beliefs are evidence-responsive mental states, their explanations for belief change should include appeals to evidence to a character acquiring new information.

Study 2 extends these findings by testing whether children can use information about prior strength to inform their *predictions* about how people will respond when encountering counterevidence to their initial beliefs. A key function of theory of mind is to predict behavior (Dennett, 1988; Ho et al., 2022; Premack & Woodruff, 1978; Wellman, 2018). In Study 2, we ask whether children predict that a character with a strong prior will be less likely to revise their belief after encountering counterevidence than a character with a weak prior. If children do not understand prior strength as interacting with evidence to produce an updated belief, prior strength should not matter for their predictions. In contrast, if children do understand the role of prior strength in belief updating, then we should expect children to predict less belief change for characters with strong priors than for those with weak priors, even when both encounter identical counterevidence.

Across both studies, we also test a crucial additional question: does children's sensitivity to prior strength extend across different types of beliefs? We examine the generalizability of children's reasoning by including four types of beliefs: a belief about the properties of a river, a belief about a creature's behavior, a belief about whether an action is wrong, and a belief about a personal preference. This selection spans both descriptive beliefs (about how the world is) and evaluative beliefs (about what is good, right, or preferred). Prior research suggests that by early elementary school, children judge that disagreement is more acceptable for evaluative judgments than for factual claims (Heiphetz et al., 2013; Wainryb et al., 2004) and that there can be more room for ambiguity or interpretation in matters of personal preference (Carpendale & Chandler, 1996). Moreover, different types of beliefs may be seen as differing in their stability and centrality to personal identity. For example, children view moral beliefs as more stable over time than preferences, and as more central to who a person is (Heiphetz, Gelman, & Young, 2017; Heiphetz, Strohminger, & Young, 2017). These distinctions could have implications for how children reason about belief revision. On the one hand, if children have intuitions about general principles of rational belief updating, sensitivity to prior belief strength should emerge across diverse belief contents. On the other hand, children may recognize that some beliefs are more amenable to revision than others, or that the same kind of evidence carries different weight in different domains. Together, these studies allow us to investigate when children integrate prior strength into their understanding of belief; specifically, when children can use this information to make modal judgments, generate predictions, and offer explanations about others' belief revision.

2. Study 1

2.1. Method

2.1.1. Participants

Participants included 100 4- to 8-year-old children (M_{age} = 6.51 years; SD = 1.44 years; range = 4.01–8.99 years; 50 boys and 50 girls) who were recruited through a university-managed database of families from a mid-sized city in the Southeastern US. This age range was chosen based on prior findings that children as young as five recognize beliefs as responsive to evidence (Confer et al., 2024; McLoughlin et al., 2021), while other work suggests that children may not understand that mental constraints like desires and norms can be overridden until around age six (Chernyak et al., 2019, 2013; Kushnir et al., 2015; Shtulman & Phillips, 2018). We aimed to capture the developmental window in which sensitivity to prior strength might emerge. Two additional children were excluded because of failure to complete the study (N = 1) and parent intervention (N = 1). An a priori G Power 3.1 analysis (Faul et al., 2007) was conducted based on a medium effect size of f^2 = 0.15 with an alpha level of 0.05, which showed that a sample size of 88 was needed to achieve 90% power for the main analyses. We preregistered a target sample size of 95 to ensure balanced representation across age groups and counterbalancing orders. Our preregistered stopping rule specified that data collection would end on the weekend of the week in which we reached 95 completed participants; this resulted in a final sample of 100 after exclusions.

Information on ethnicity collected from parental questionnaires showed that the participants were 76.8%White, 1% Black, 8.1% Asian, 4% Hispanic or Latino, and 10.1% identified their child with multiple races or ethnicities (99 of the 100 parents completed this question). Self-reported socialeconomic status revealed that 59.2% of the primary caregiver had a graduate/professional school degree, 37.8% a college diploma, 3% completed high school or lower. For the participants' other primary caregiver, 53% had a graduate/professional school degree, 32.7% a college diploma and 14.3% had a high school diploma or lower. Parent-reported income was primarily middle to upper class, with 61.2% of the parents reporting an annual income greater than $100,000 per year, 26.5% reported an income between $50,000 and $100,000, and 12.3% reported an annual income lower than $50,000 (98 of the 100 parents completed these questions). All participants were tested in online sessions via videoconferencing with a live experimenter, or at the museum.

2.1.2. Procedure

Children evaluated four sets of characters with beliefs of different kinds; each belief involved novel words to avoid children's own beliefs influencing their judgments. Two of the beliefs were descriptive: one about social behavior (“Flurpies dance before bedtime”), and one a fact about a river (“The Luzak river is the longest river”), and two were evaluative: one about moral norms (“Daxing is wrong”) and one about personal preferences (“Pelling tastes yummy”). The presentation order of belief types was counterbalanced using a Latin-square design.

In each trial, children were introduced to two characters who held the same belief content, but differed in terms of strength: one character believed the target claim a lot (strong initial belief), and the other believed it only a little (weak initial belief). For example, in the moral norms trial, children were introduced to two characters who believed that “daxing”—an action one person could do to another—was wrong. The experimenter then described each character's belief strength explicitly (e.g., “[Sarah/Justine] believes that daxing is wrong. She believes it [a lot/a little bit]; she is [very sure/only a little sure] that daxing is wrong. Whenever someone asks her what she believes, she says ‘[I'm very sure that daxing is wrong/Hmmm…I'm only a little sure that daxing is wrong]’, see Fig. 1).

After presenting both characters' beliefs, the experimenter first asked *comprehension check* questions to confirm that children remembered the strength of each character's belief (e.g., “How much does [Sarah/Justine] believe that [daxing is wrong]? Does she believe it a lot or a little?”). Children received corrective feedback if they could not remember the strength of each character's beliefs. Next, the experimenter asked a *belief change possibility* question for each character, for example: “Can [Sarah/Justine] change her belief and believe that [daxing is *not* wrong]?” Responses were scored as 1 if children judged belief change to be possible and 0 if they judged that the character could not change their beliefs. Finally, the experimenter asked a *belief change difficulty* question for each character, for example: “Is it hard or easy for [Sarah/Justine] to change her belief and believe that [daxing is not wrong]?” Based on their initial response, children were asked whether it was “a little [hard/easy]” or “really [hard/easy],” and responses were converted into a 4-point scale (really easy to change belief = 1, really hard to change belief = 4).

After completing all four trials, participants completed a *belief change mechanism task* where they provided explanations for how a character might change their belief. The experimenter reminded children of one character from each trial (either the character with the strongly-held or weakly-held beliefs, counterbalanced across participants), for example: “Remember Sarah? She was the one who believes that *daxing* is wrong and she believes it a lot; she is very sure that *daxing* is wrong.”). The experimenter then described the character as changing their belief, and asked an open-ended question about how belief change happened (e.g., “Let's say that one day, Sarah changed what she believes, and she now believes that [daxing is not wrong].What do you think happened? How did [Sarah]change her belief?”) Children's responses were transcribed and coded from video recordings. If children did not respond spontaneously after being prompted twice, the experimenter provided them with three options, asking if the character saw something, heard someone say something, or decided to change their belief.

2.1.3. Transparency statement

We preregistered this study on AsPredicted (#133945) after collecting pilot data from 26 participants to confirm that children understood the study questions and to obtain an effect size estimate for power analysis. No procedural changes were made following piloting, and no analyses were conducted on pilot data prior to preregistration. Pilot participants were included in the final sample because the procedure remained unchanged; robustness analyses excluding pilot participants yield the same pattern of results (see Supplemental Materials). The preregistration specified our primary hypotheses, sample size and stopping rule, key predictors, primary outcome measures, and the use of mixed-effects models with participant ID as a random effect. The preregistration also specified that interaction terms would be tested and retained if they improved model fit, and that analyses of belief type/domain would be exploratory.[2] Model-building procedures (e.g., likelihood ratio tests for comparing models) and follow-up analyses (e.g., simple slopes) were not fully specified in the preregistration.

2.2. Results

We report results for three measures: (1) belief change possibility judgments (binary responses; 1 = belief can change, 0 = belief cannot change), (2) belief change difficulty judgments (4-point scale, 1 = very easy to 4 = very hard), and (3) open-ended explanations for mechanisms of belief change. For the belief change possibility and belief change difficulty measures, we ran multilevel models using the lme4 package

[2] Our preregistration specified combining the Change Possibility and Change Difficulty responses into a single 5-point score, though it also allowed for analyzing them separately. We report the separate analyses here because possibility and difficulty are conceptually distinct and use different response scales. The preregistered combined-score analysis is reported in the Supplemental Materials and yields a consistent pattern of results.

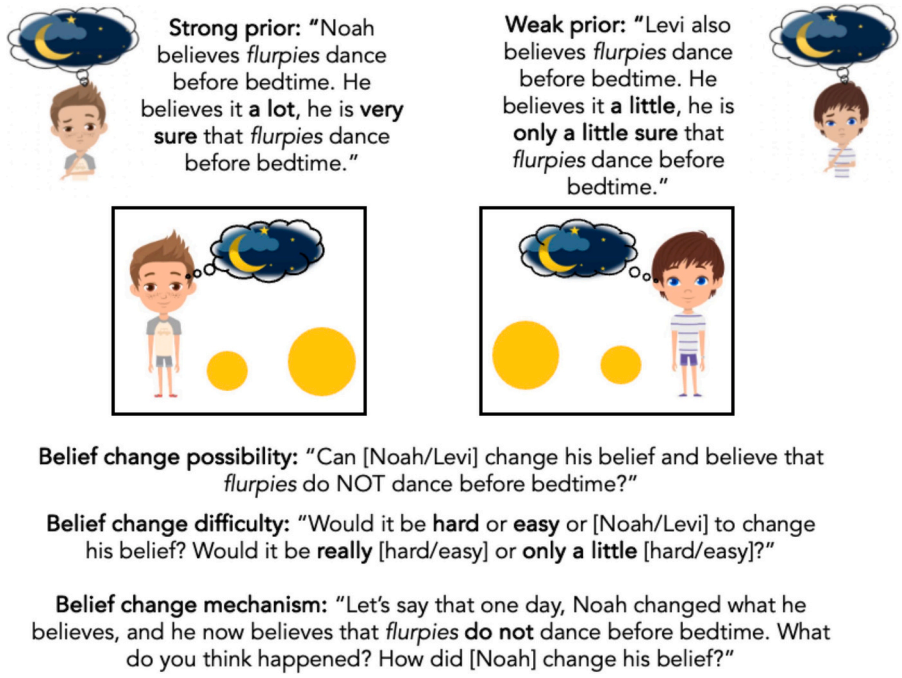


Fig. 1. Study 1 procedure overview.

(Bates et al., 2015) in R. Age (continuous, centered), belief strength (strong initial belief, weak initial belief), and belief type were included as fixed effects; participant was included as a random intercept. Interaction terms were added sequentially, and we used likelihood ratio tests to evaluate whether each interaction significantly improved model fit. Belief type was included as an exploratory predictor but did not show significant effects and was excluded from final models. Preliminary analyses confirmed no significant trial order effects (all $ps > 0.44$).

2.2.1. *Belief change possibility question ("Can [character] change her belief?")*

We first examined children's judgments about whether belief change is possible, based on their responses to the question "Can [character] change her belief and believe that [X]?" Across all trials, children judged that beliefs could change 85.5% of the time. This proportion was higher for weakly held beliefs ($M = 91.7\%$) than for strongly held beliefs ($M = 80.6\%$). Before examining predictors of children's belief change judgments, we first assessed whether their responses differed from chance (50%) using Benjamini-Hochberg (BH) adjusted binomial tests. Across all age groups, children judged that beliefs—both weakly and strongly held—could change at rates significantly above chance ($ps < 0.001$). For weakly-held beliefs, the proportion of belief change judgments increased from 76.5% at age 4 to 100% at age 8. For strongly-held beliefs, this proportion also exceeded chance across all age groups, ranging from 70.7% at age 4 to 90.5% at age 8. These findings suggest that even the youngest children in our sample were more likely to view beliefs as possible to be updated.

Next, we tested the effects of age and prior belief strength on children's belief change possibility judgments (binary responses; 1 = belief can change, 0 = belief cannot change) using mixed-effects logistic regression models. We found a significant main effect of age ($b = 1.56$, $SE = 0.41$, OR = 4.78, 95% CI [2.15, 10.6], $p = .001$), indicating that older children were more likely to judge beliefs as changeable. There was also a significant main effect of belief strength, such that strongly held prior beliefs were less likely to be judged as changeable than weakly held ones ($b = -2.56$, $SE = 0.46$, OR = 0.08, 95% CI [0.03, 0.19], $p < .0001$). In addition, we found that a significant interaction between age and belief strength ($b = -1.09$, $SE = 0.30$, OR = 0.34, 95% CI = [0.19, 0.6], $p = .0003$), indicating that with age, children became more likely to judge that weakly held beliefs can change compared to strongly-held beliefs (see Fig. 2).

To further unpack the interaction between age and belief strength, we conducted a simple slopes analysis. For weakly-held beliefs, we found a significant main effect of age ($b = 1.56$, $SE = 0.41$, OR = 4.78, 95% CI [2.15, 10.6], $p = .0001$): older children were more likely to judge that weakly-held beliefs can change. In contrast, there was no significant effect of age for the strongly-held belief judgments ($b = 0.47$, $SE = 0.31$, OR = 1.6, 95% CI [0.87, 2.95], $p = .13$), suggesting that judgments about strongly-held beliefs remained relatively stable across the age range tested.

2.2.2. *Belief change difficulty question ("Is it hard or easy for [character] to change her belief?")*

We next examined children's judgments about how difficult belief change would be (rated on a 4-point scale, very easy = 1, very hard =4). On average, children rated belief change as moderately difficult ($M = 2.42$, $SD = 1.13$). Children rated strongly held beliefs as more difficult to change ($M = 2.85$, $SD = 1.1$) than weakly held beliefs ($M = 1.98$, $SD = 0.98$). We analyzed these responses using linear mixed models (LMMs). The best-fitting model included age (continuous, centered), belief strength (strong initial belief, weak initial belief), and their interaction as fixed effects, and participant as a random intercept. There was a significant main effect of belief strength: participants rated strongly held beliefs as more difficult to change than weakly held ones ($b = 0.87$, $SE = 0.06$, 95% CI [0.76, 0.99], $p < .001$). There was also a significant main effect of age ($b = -0.11$, $SE = 0.05$, 95% CI [−0.21, −0.01], $p = .03$). These effects were qualified by a significant age by belief strength interaction ($b = 0.27$, $SE = 0.04$, 95% CI = [0.18, 0.35], $p < .001$): with age, children increasingly distinguished between strong and weak beliefs in their difficulty judgments (see Fig. 3). To unpack the interaction, we conducted a simple slopes analysis. For strongly-held belief judgments, we found a significant positive effect of age ($b = 0.16$, $SE = 0.05$, $p = .002$), indicating that with age, children judged strongly-held beliefs to be increasingly difficult to change. For weakly-held belief judgments, we found a significant negative effect of age ($b = -0.11$, $SE = 0.05$, $p = .03$), indicating that with age, children judged weakly-held beliefs to be

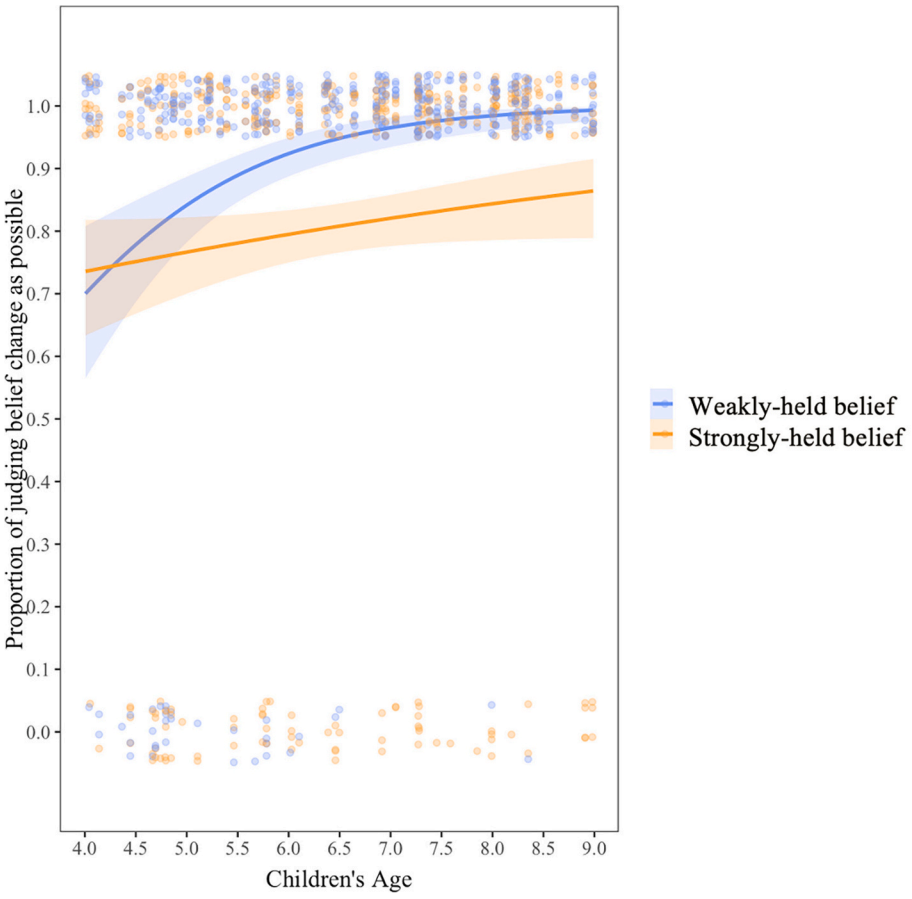


Fig. 2. Study 1: children's belief change possibility judgments as a function of age and belief strength.
Note. Points represent individual raw data (jittered for visibility). Shaded ribbons represent 95% confidence intervals around the regression lines.

increasingly easy to change.

2.2.3. Change mechanism task (e.g., "How did [character] change her belief?")

Finally, we examined children's explanations for how a character might change their belief, based on their responses to the question "What do you think happened? How did [character] change her belief?" For the present analyses, we focused on the open-ended responses that children provided before receiving any forced-choice prompt. This decision was based on the reasoning that open-ended explanations offer a more unprompted measure of how children spontaneously reason about belief change. While the forced-choice options were useful for eliciting responses from participants who did not initially answer, they may have influenced children's reasoning by introducing specific mechanisms of belief change. Descriptive analyses combining open-ended and forced-choice responses yielded similar overall patterns. Children's open-ended explanations as to why a character's belief changed were transcribed from videotaped sessions and coded by two coders. Reliability coding was then calculated on the basis of 25% of the justifications. Interrater reliability was high (92%). Discrepancies were resolved through discussion.

Children's explanations for belief change were assigned to one of five categories: (1) *First-hand evidence* (2) *Second-hand evidence*; (3) *Internal cognition*; (4) *Desire to change*; and (5) *Irrelevant or no response* (see definitions and examples of the categories in Table 1).

The most common explanations across all belief types were *first-hand evidence*, which was cited most often for the personal preference belief (62%) and least often for the moral belief (29%). Conversely, second-hand evidence was more frequent for moral norm (21%) and factual beliefs (24%), but comparatively less common for the personal preference belief (10%). Notably, internal non-evidential explanations were infrequent overall (4% - 9%, see Table 2 for proportions of each category by belief type).

In exploratory analyses, we conducted separate binary logistic generalized linear mixed-effects models (GLMMs) for each explanation category. Each model predicted whether a given explanation type was used (1) or not (0), with belief domain and age as fixed effects and participant as a random intercept. The likelihood of children providing *first-hand evidence* explanations varied significantly by belief domain, χ^2 (3) = 25.89, $p < .001$. Similarly, the use of *second-hand evidence* explanations differed across domains, χ^2 (3) = 12.74, $p = .005$. Age was not a significant predictor of evidence-based explanations in either model (all *ps* > 0.51). Explanations involving a *desire to change* or internal cognition did not vary by age (*ps* > 0.34). or domain (*ps* >. 053). Post-hoc pairwise comparisons (Benjamini-Hochberg adjusted) confirmed that children were most likely to provide *first-hand evidence* explanations in the personal preference domain, which significantly exceeded all other domains (all *ps* < 0.007). For second-hand evidence, children were more likely to reference this explanation type for moral and factual beliefs compared to personal preferences (*ps* = 0.01). These findings may suggest that children were most likely to rely on direct, perceptual evidence when explaining belief change for personal taste preferences, while their reasoning about moral norms may involve both first- and second-hand sources of information.

2.3. Discussion

Study 1 examined whether 4- to 8-year-old children make assumptions that other people are epistemically rational; that is, whether young children assume that the strength of prior beliefs constrains a person's

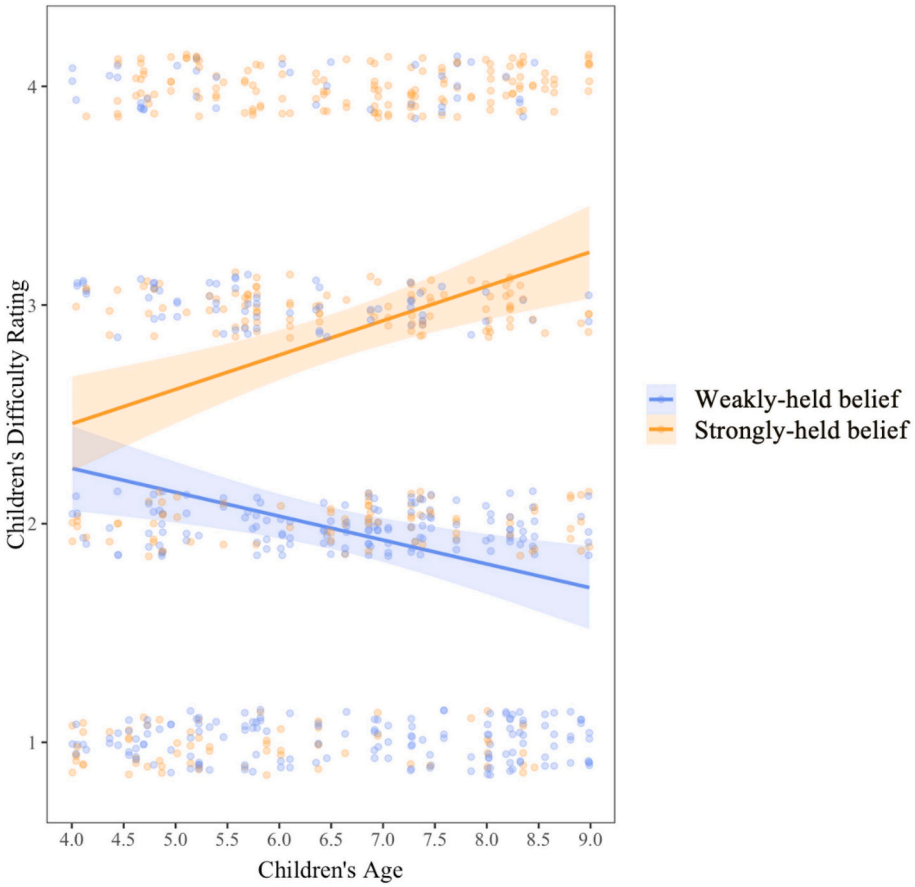


Fig. 3. Study 1: Children's belief change difficulty judgments as a function of age and belief strength.
Note. Points represent individual raw data (jittered for visibility). Shaded ribbons represent 95% confidence intervals around the regression lines.

Table 1
Study 1: children's justification categories for belief change.

Response Category	Description	Example
First-hand evidence	Explaining that belief change happens after first-hand observations, exploration or other direct experience.	"He saw the flurpies *not* dance before bedtime."
Second-hand evidence	Explaining that belief change happens after hearing opposite beliefs or counter-evidence from adults, peers or other second-hand sources (e.g., books).	"She heard that other kids found out the truth so she asked what happened and they told her the truth."
Internal cognition	Explaining that belief change happens after thinking, reasoning or recalling something through memory.	"Maybe he thinks about it"
Desire to change	Reasoning that the character has a desire to change their mind.	"He just wanted to change his mind."
Irrelevant	Giving other justifications that are not listed above or not related to the vignette.	"I don't know. He changed it to taco soup."

Table 2
Study 1: frequencies of children' justification categories for belief change by belief type.

Response category	Social behavior (flurpies)	Fact (river)	Moral (daxing)	Personal (pelling)
First-hand evidence	44.23%	40.38%	28.71%	61.76%
Second-hand evidence	14.42%	24.04%	20.79%	9.80%
Internal cognition	7.69%	7.69%	7.92%	4.90%
Desire to change	3.85%	6.73%	8.91%	2.94%
Irrelevant	26.92%	24.04%	33.66%	20.59%

ability to change their beliefs. Our results show that children across these ages judged that beliefs can change, even strongly held ones; but starting around age five, children judged strongly held beliefs to be more difficult to change than weakly-held beliefs. With age, children increasingly differentiated between strongly- and weakly-held beliefs in their possibility and difficulty judgments. Importantly, this same pattern emerged across all kinds of belief, suggesting that children understand this as a domain-general feature of how beliefs work. These results provide clear evidence that by age five, children treat prior belief strength as a constraint on belief revision and suggest that children's theory of mind includes an intuitive understanding of belief change that aligns with principles of rational belief revision. This extends previous research revealing children's sensitivity to mental constraints (Confer et al., 2024; Kushnir et al., 2015), suggesting that children treat strong priors as an epistemic constraint that may require proportionally stronger counterevidence to overcome—a pattern that aligns with the Bayesian framework of rational belief updating (Gärdenfors & Makinson, 1988; Titelbaum, 2022). Notably, although children did not distinguish between the revisability of strong versus weak beliefs until age five, even four-year-olds judged belief revision to be possible at above chance levels, even for strongly held beliefs. Our findings suggest that even if young children judge changing a strong belief to be difficult–and therefore, potentially unlikely–they still see it as *possible*.

Children's judgments about belief change difficulty showed robust developmental effects. With age, children were more likely to judge that

strongly held beliefs are more difficult to change than weak ones. Importantly, this was not just a matter of older children seeing belief change as more effortful overall. Rather, with age, children viewed strongly held beliefs as increasingly difficult to change, while weakly held beliefs were perceived as increasingly easy to change. This developmental pattern suggests that children's early understanding of belief revision may incorporate epistemic effort.

Finally, the open-ended explanations suggest that children often think that belief change is driven by the acquisition of new information, consistent with normative theories of rational belief. Across belief domains, the majority of children's responses appealed to first-hand or second-hand evidence as the mechanisms for belief revision. Very few children attributed belief change to internal desire. This pattern suggests that children do not see belief change as arbitrary or internally driven; rather, they view belief change as rooted in external evidence. Evidence-responsiveness is an essential feature of the nature of belief (Feldman & Conee, 1985; Williams, 1973); the fact that children appeal to evidence as a primary mechanism for belief change is an important piece of evidence supporting the claim that they have an emerging understanding of how beliefs work.

3. Study 2

In Study 1 we found evidence in support of the idea that children's understanding of belief change incorporates expectations of epistemic rationality. Our results revealed that by age 5, children judge strongly held beliefs as less possible and more difficult to change than weakly held ones, and that children's explanations for belief change appeal to evidence. This offers affirmative evidence for two key questions: first, whether children are sensitive to the idea of graded prior beliefs, and second, whether they assume that prior belief strength constrains belief updating. Study 2 builds on these findings by testing whether children integrate prior strength into their predictions about belief revision. Specifically, we ask whether children use information about prior belief strength to predict how someone will respond to counterevidence.

As in Study 1, we presented children with pairs of characters who held the same belief but differed in belief strength. This time, however, both characters encountered counterevidence to their original belief: a group of peers who unanimously expressed the opposing view. We chose peer testimony as our operationalization of counterevidence for several reasons. First, presenting testimonial evidence allowed us to apply structurally parallel counterevidence across all belief domains, since direct observation might be applicable to some beliefs (e.g., whether *Flurpies* dance) but less straightforwardly to others (e.g., whether *daxing* is wrong). Second, testimony is a primary and powerful source of evidence for young children (Gelman, 2009; Harris et al., 2018), and prior work shows that children treat peer and adult consensus as reliable sources of second-hand evidence (Corriveau et al., 2009; Haun & Tomasello, 2011) and moderate their belief updating based on both their own prior confidence and the perceived strength of the testimony (Langenhoff et al., 2023). Third, testimonial counterevidence is probabilistic and socially mediated: children can reasonably treat peer consensus as informative without treating it as determinative. This makes testimonial counterevidence well-suited for testing whether children's predictions track the influence of prior strength. Our primary prediction was that children would use prior strength to predict belief change: characters with strong priors should be judged less likely to revise their beliefs than those with weak priors, even when both encounter the same counterevidence. As in Study 1, we tested this prediction across different belief domains to examine the generality of children's understanding of belief updating. A secondary, exploratory question concerned whether children's predictions might differ across belief types. Within a rational belief-updating framework, the impact of evidence depends not only on priors but also on the perceived strength of the evidence which could conceivably differ by domain. For instance, children might view peer consensus as informative for descriptive claims (e.g., "Flurpies dance before bedtime") than for a subjective personal preference (e.g., "Pelling is yummy"), where individual variation is expected (Carpendale & Chandler, 1996). Similarly, peer testimony may carry less epistemic weight for moral beliefs if children view such beliefs as requiring justification from authoritative sources rather than peer agreement (Laupa & Turiel, 1986).

3.1. Method

3.1.1. Participants

We recruited 99 4- to 8-year-old children using the same methods as Study 1 (M_{age} = 6.48 years, range = 4.03–8.99 years, 51 boys and 48 girls; 64.6% White, 3% Black, 8% Asian, and 24.4% were identified with multiple races). Three additional children were excluded from the final analyses because of technical issues (N = 1) or experimenter error (N = 2). Self-reported education status of parents was high: 67.7% of first primary caregivers had a graduate/professional school degree, 30.3% a college diploma, and 2% completed high school or lower. For the participants' other primary caregiver, 55.1% had a graduate/professional school degree, 39.8% a college diploma and 5.1% had a high school diploma or lower. 87.5% of the parents reporting an annual income greater than $100,000 per year, 11.5% reported an income between $50,000 and $100,000, and 1% reported an annual income lower than $50,000. 79.8% of participants were tested in online sessions via videoconferencing with a live experimenter, and 20.2% of participants were tested in a museum (there were no significant differences between testing settings across children's responses).

3.1.2. Procedure

As in Study 1, children were presented with two characters who initially held the same belief content, but differed in strength: one character had a strong initial belief, the other a weak initial belief. Children again completed four trials with different kinds of belief. We used the same four beliefs from Study 1, with one modification: "The Luzak River is the longest river on the planet" was replaced with "The Luzak River has the most fish of any river on the planet," to account for some children challenging this belief. The order of belief types was counterbalanced across participants using a Latin-square design.

In each trial, after learning about the characters' initial beliefs, children again answered comprehension check questions to confirm their understanding of each character's initial belief strength. Next, children were told that both characters overheard a group of five people each saying that the target belief was false (see Fig. 4). For example, children heard: "One day, a group of five people are talking about daxing. Sarah and Justine both overhear them talking. They hear all five people say what they think, and each person says, 'I think daxing is *NOT* wrong; I think daxing is okay.'"

Children then answered two questions about each character. First, after being reminded of the strength of the characters' initial belief and the opposing belief from the group, children answered a *belief change prediction question* assessing whether the character would maintain or change their beliefs after encountering counterevidence. For example: "If before, [Sarah believes that daxing is wrong a lot and is really sure that daxing is wrong, and now hear all five people in the group say that daxing is okay], would [Sarah] change her belief and believe that [daxing is not wrong], or keep her belief that [daxing is wrong]?" This was followed by an *updated belief strength* question assessing what children thought the character's updated belief would be after they encountered the counterevidence. For example: "So you think [Sarah] would [change her belief and now believe that daxing is NOT wrong/keep her belief and still believe that daxing is wrong]. How much would she believe this? A little or a lot?" Children's responses were converted a four-point *updated belief score*, in which higher scores indicate stronger belief in the negated original belief (new belief a lot = 4; old belief a lot = 1).

In addition, on the final trial, children completed an exploratory

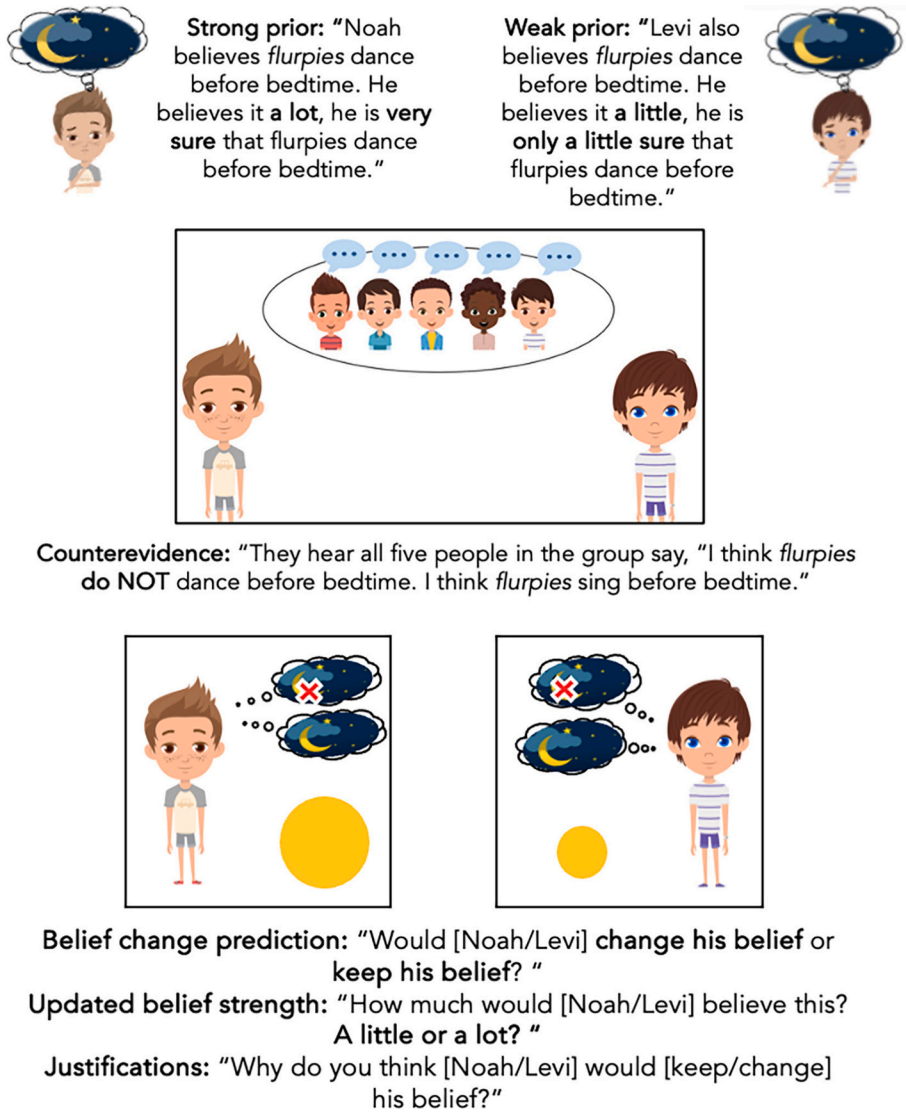


Fig. 4. Study 2 procedure overview.

measure in which they provided open-ended justifications for why each character would or would not revise their belief. For example, children were asked: "Why do you think [Sarah] would [change/keep] her belief?" Children's responses were recorded.

3.1.3. *Transparency statement*

We preregistered this study on AsPredicted (#176514) after collecting pilot data from 27 participants. As in Study 1, no procedural changes were made and no analyses were conducted prior to preregistration. All key effects replicate when excluding pilot participants collected before preregistration (see Supplemental Materials). The preregistration specified our primary hypothesis, sample size and stopping rule, key predictors, primary outcome measures, and the use of mixed-effects models. Analyses of belief type/domain were preregistered as exploratory. Model-building procedures and follow-up analyses were not fully specified in the preregistration. Study materials, data and preregistrations for both studies are available on the Open Science Framework: https://osf.io/ng5dr/. Both studies were approved by Duke University IRB.

3.2. *Results*

We report results for three measures: (1) belief change predictions (binary responses: 1 = change, 0 = keep), (2) updated belief strength scores (4-point scale, 1 = strongly endorses original belief to 4 = strongly endorses new belief), and (3) open-ended justifications for updated belief judgments. For the first two measures, we used mixed-effects models with age (continuous, centered), belief strength (strong vs. weak), and belief type (moral, social behavior, factual, personal) as predictors, with participant as a random intercept. Preliminary analyses confirmed no significant order effects (all $ps > 0.45$).

3.2.1. *Belief change prediction question ("Would [character] change her belief or keep her belief?")*

Before examining the primary outcome (updated belief scores), we report children's binary judgments about whether characters would change or keep their beliefs after encountering counterevidence (e.g., "Would [Sarah] change her belief and believe that *daxing* is not wrong, or keep her belief that *daxing* is wrong?"). Overall, children predicted belief change on 40.3% of trials. This proportion was higher for weakly held beliefs (M = 50.3%) than for strongly held beliefs (M = 30.3%). Children's binary responses about whether the character would change (coded as 1) or keep (coded as 0) their belief were analyzed using mixed-effects logistic regression models. The final model included fixed effects of age, belief strength, belief type, and interactions of age with both belief Strength and belief type, with a random intercept for participant

ID.

We found a significant main effect of belief strength ($b = -1.34$, $SE = 0.20$, OR = 0.26, 95% CI = [0.18, 0.39], $p < .001$): children predicted lower likelihood of change for strongly held beliefs than weakly held ones. This was qualified by a significant age by belief strength interaction ($b = -0.48$, $SE = 0.13$, OR = 0.62, 95% CI = [0.48, 0.81], $p = .0004$): indicating that with age, children were more likely to distinguish between strong and weak beliefs in their predictions (see Fig. 5). We followed up on this interaction by assessing the effect of age separately for strong and weak prior beliefs using simple slope analyses. For strongly held beliefs, there was a significant negative effect of age ($b = -0.50$, $SE = 0.16$, OR = 0.61, 95% CI [0.44, 0.83], $p = .002$): older children were less likely to predict belief revision for strongly held beliefs compared to younger children. In contrast, age did not significantly predict belief change judgments for weakly held beliefs ($b = -0.02$, $SE = 0.15$, OR = 0.98, 95% CI [0.72, 1.32], $p = .89$).

We also found a significant interaction between age and belief type ($\chi^2(3) = 10.85$, $p = .01$), indicating that age-related changes in children's judgments varied across different domains of beliefs (see Fig. 6). To unpack this interaction, we conducted simple slope analyses for age within each belief type (averaged across belief strength). We found that age negatively predicted belief change judgments in the moral domain ($b = -0.53$, $SE = 0.19$, OR = 0.59, 95% CI [0.41, 0.84], $p = .004$), suggesting older children were less likely to expect moral beliefs to change. A similar, marginally significant trend was observed for social behavior beliefs ($b = -0.32$, $SE = 0.18$, OR = 0.73, 95% CI [0.51, 1.04], $p = .08$). In contrast, Age was not significantly related to judgments for personal (OR = 0.78, 95% CI [0.54, 1.10], *ns*) or factual (OR = 1.06, 95% CI [0.75, 1.51], *ns*) beliefs. In addition, pairwise comparisons of these age-related trends revealed that the decline in belief change expectations was significantly steeper for moral beliefs compared to factual beliefs ($p = .009$). There were no significant differences between other pairs of trials (all $p > .12$).

3.2.2. *Updated belief strength question*

We next examined children's predictions about the strength of characters' beliefs after encountering counterevidence (e.g., "How much would she believe this? A lot or a little?"). Responses were scored on a 4-point scale (endorse original belief a lot = 1, endorse new belief a lot = 4). Importantly, if children think the character did not change their mind, their rating should reflect how strongly the character kept the original belief. For example, if children think a character with a strong prior would maintain their belief, they would receive a score of 1 (still strongly endorsing the original belief), whereas a character with a weak prior who maintained their belief would receive a score of 2 (still somewhat endorsing the original belief). On average, updated belief scores were moderate ($M = 2.27$, $SD = 1.11$), with higher scores (indicating more predicted change) for weakly held beliefs ($M = 2.59$, $SD = 0.98$) than for strongly held beliefs ($M = 1.94$, $SD = 1.14$).

These responses were analyzed using linear mixed-effects models with random intercepts for participant ID. The final, best-fitting model included fixed effects of age, initial belief strength, belief type, and the interactions of age with both initial belief strength and belief type.

We found a significant interaction between age and initial belief strength ($b = -0.15$, $SE = 0.04$, 95% CI [−0.24, −0.06], $p = .0008$), indicating that children's predictions about updated belief strength e depended on both the strength of the character's initial belief and the child's age (see Fig. 7). To explore this interaction, we conducted simple slope analyses examining the effect of age separately for strong and weak prior beliefs. For strongly-held beliefs, there was a significant

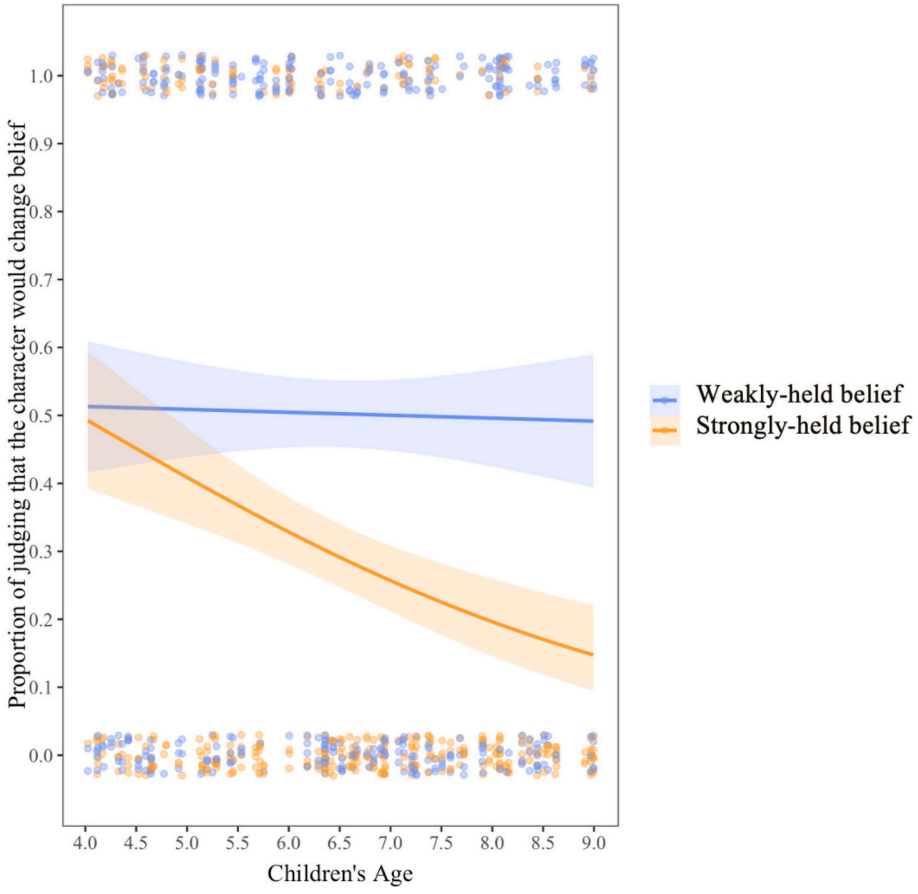


Fig. 5. Study 2: children's belief change prediction judgments as a function of age and belief strength.
Note. Points represent individual raw data (jittered for visibility). Shaded ribbons represent 95% confidence intervals around the regression lines.

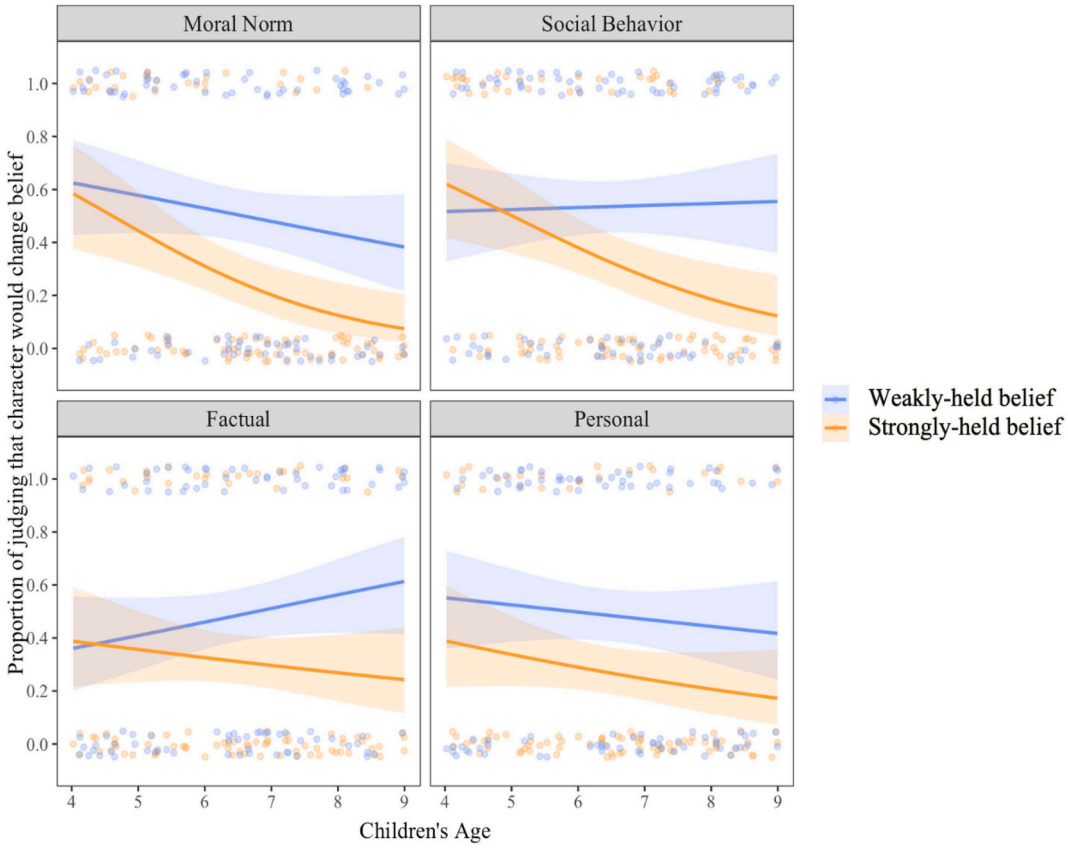


Fig. 6. Study 2: Children's belief change prediction judgments separated by belief type.
Note. Points represent individual raw data (jittered for visibility). Shaded ribbons represent 95% confidence intervals around the regression lines.

negative effect of age ($b = -0.15$, $SE = 0.05$, 95% CI [−0.25, −0.05], $p = .004$), revealing that older children predicted less change in strong initial beliefs compared to younger children. In contrast, for weakly-held initial beliefs, age did not significantly predict judgments ($b = 0.001$, $SE = 0.05$, 95% CI [−0.10, 0.10], *ns*).

In exploratory analyses, we also found a significant interaction between age and belief type ($\chi^2(3) = 9.94$, $p = .02$), indicating that age-related changes varied by belief domain. Follow-up simple slope analyses showed that age negatively predicted updated belief judgments only for the moral norm belief trial ("daxing is wrong;" $b = -0.17$, $SE = 0.06$, 95% CI [−0.29, −0.06], $p = .004$), indicating that older children increasingly expected moral beliefs to be resistant to change. Age effects for other beliefs-types were not significant (all $ps > 0.10$). Pairwise comparisons of these age-related slopes across trial types (Benjamini–Hochberg adjusted) revealed that the decline with age was significantly steeper for moral beliefs compared to factual beliefs ($p = .01$). No other pairwise comparisons between belief types were statistically significant (all $ps > 0.21$).

3.2.3. *Justifications for updated belief judgments*

As an exploratory measure collected on the final trial, we examined children's open-ended justifications for their predictions. When asked to explain why they thought the character would keep or change their belief, children's justifications could be assigned to one of five categories: (1) *Reasoning about prior belief;* (2) *Reasoning about counter-evidence*; (3) *Personal Opinion;* (4) *Child's Own Beliefs*; and (5) *Irrelevant or no response* (see definitions and examples of the categories in Table 3). Reliability coding was then calculated on the basis of 25% of the justifications. Interrater reliability was high (92.9%). Discrepancies were resolved through discussion.

We found that children's reasoning focused on *Prior beliefs* the most ($n = 47$; 34.3%) followed by reasoning that it is acceptable for the target characters to have their *Own opinions* ($n = 35$; 25.5%); reasoning about *Counterevidence* ($n = 31$; 22.6%) and children's own beliefs ($n = 24$; 17.5%). In general, these exploratory findings suggest that children took both the characters' prior beliefs and the counterevidence into account in their judgments. To examine whether children's justification patterns varied with age and across prior strength, we conducted separate binary logistic GLMMs for each justification category as exploratory analyses. Each model included prior strength (strong vs. weak) and age (continuous, centered) as fixed effects, with participant as a random intercept. The analyses revealed no significant effects of prior strength (all $ps > 0.16$) or age (all $ps > 0.45$) on any of the justification categories. Thus, children across the age range drew on similar types of reasoning when justifying their judgments.

3.3. *Discussion*

In Study 2, we explored whether children use information about a character's initial belief strength to predict how that character would update their beliefs in response to counterevidence. Overall rates of predicted change were around 50%—but most importantly for our primary research question, predicted change depended on prior strength. By around age five, children predicted that characters with strong initial beliefs would be more likely to keep their original beliefs than characters with weak initial beliefs. Furthermore, when asked to estimate the strength of the character's belief after encountering opposing evidence, children's predictions reflected sensitivity to the strength of the prior belief. These findings provide additional evidence that young children expect belief revision to be epistemically rational: they think that belief change is more likely for weak beliefs than strong ones, and they use information about prior strength to make predictions about what an

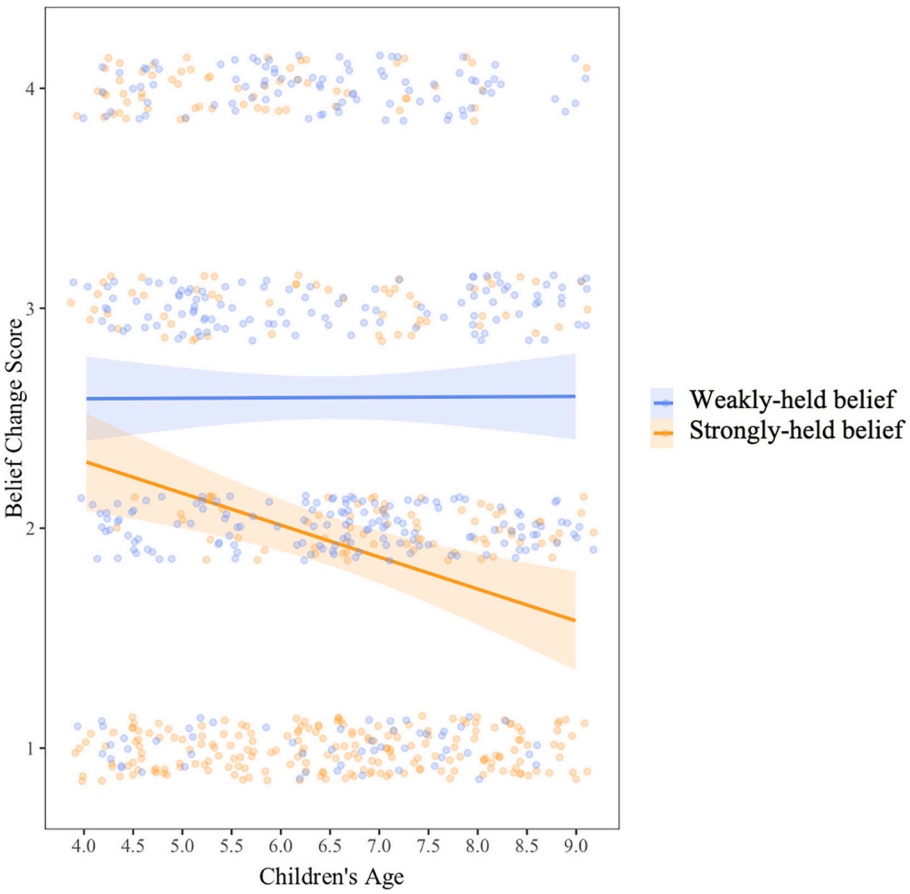


Fig. 7. Study 2: children's updated belief strength judgments as a function of age and initial belief strength.
Note. Points represent individual raw data (jittered for visibility). Shaded ribbons represent 95% confidence intervals around the regression lines.

Table 3
Study 2 children's justification categories for the likelihood of belief change.

Category Name	Description	Example	%
Reasoning about prior belief	Focusing on the character's prior belief or commenting on the strength of the character's initial belief	"Probably because he only believed it a little bit."	34.3%
Reasoning about counter-evidence	Focusing on the counter-evidence from the group of peers	"He changed his mind just because a lot of people said it didn't taste good."	22.6%
Personal opinion	Reasoning about how it is acceptable for the character to have different opinions or preferences than others	"Because it's her opinion and other people have other opinions."	25.5%
Child's own belief	Making predictions based on how likely the participant think the prior belief is valid.	"Because I just think daxing might be a little wrong."	17.5%
Irrelevant	Giving other justifications that are not listed above or not related to the vignette.	"I don't know. He changed it to taco soup."	32.8%

agent's updated belief will be after encountering novel counterevidence. Children's open-ended justifications also explicitly referenced prior belief strength or the counterevidence, again reflecting that children recognized the relevance of both initial belief strength and new information in determining the likelihood of change.

As in Study 1, children's sensitivity to prior belief strength emerged across a range of belief types. This suggests that regardless of the content of the belief, children reason that the strength with which an initial belief is held affects how likely it might be affected by counterevidence. At the same time, our findings also reveal some domain-specific differences in how children predicted characters would respond to the counterevidence presented in this study: specifically, older children became less likely to predict belief change in the moral norm trial ("daxing is wrong") compared to other trials. These exploratory effects may suggest that older children evaluate the evidential weight of peer testimony differently across domains. Peer consensus may be seen as compelling counterevidence for factual claims, but less authoritative for moral beliefs (Sebastián-Enesco et al., 2020).

An additional consideration is whether children's predictions in Study 2 reflect expectations about genuine belief revision or expectations about conformity to a peer majority. Research has shown that young children do conform to peer or adult majorities, sometimes even against their own perceptual judgments (Corriveau & Harris, 2010; Haun & Tomasello, 2011). If children were predicting that characters would capitulate to social pressure rather than revise their beliefs, this would represent a different psychological process than the epistemic updating we aimed to study. Several features of our findings, however, suggest that children's predictions reflect reasoning about epistemic processes rather than conformity expectations. First, if children expected characters to conform to peer pressure regardless of their epistemic states, we would not observe the robust effect of prior strength: social pressure from a unanimous group was constant across conditions, yet children predicted less belief change for strong-prior characters. Second, children's open-ended justifications often referenced the character's prior belief strength or the informational content of the testimony, rather than social pressure or the desire to fit in. Third, the domain-specific patterns we observed are difficult to explain on a pure conformity account. That said, future research could more directly distinguish conformity expectations from belief revision predictions by

separately asking whether a character would publicly agree with the group versus privately believing the group is right. If conformity expectations drive children's predictions, they might predict public agreement regardless of prior strength while reserving predictions about private belief change for weak-prior characters (Haun & Tomasello, 2011).

Interestingly, some children also appealed to the idea that belief change is optional: that it is up to a person whether they want to change their beliefs, and that they are permitted to believe what they want. Specifically, they reasoned in their justifications that it is acceptable for someone to hold a different belief or opinion from others. These responses suggest that children are developing not only an understanding of how beliefs change, but also a sense that belief revision might be normatively flexible: that sometimes, there is latitude in what an agent can permissibly believe, and this may vary based on the type of the belief in question. Although this type of reasoning did not differ significantly across belief domains in our study, future research could investigate whether this tolerance for different patterns of belief-updating is more prevalent in evaluative domains like morality or personal preferences compared to factual domains, potentially consistent with prior findings that children think multiple people can be right in some domains but not others (Heiphetz et al., 2013).

4. General discussion

Across two studies, we investigated an underexplored aspect of children's intuitive theory of mind, asking whether children expect others to be epistemically rational. Specifically, we investigated the role of prior strength in children's judgments, predictions, and explanations of belief revision. Our findings show that by age five, children treat prior belief strength as a meaningful epistemic constraint on belief revision. They consider it in judgments about the possibility and difficulty of belief change (Study 1) as well as in their predictions about how a character would respond to counterevidence (Study 2). Moreover, this pattern emerged across a range of belief types, suggesting that children treat epistemic rationality as a constraint on belief revision that is domain-general. These findings extend prior work showing that children are sensitive to the certainty with which one holds a belief (Sabbagh & Baldwin, 2001), understand epistemic constraints on free action (Chernyak et al., 2013; Kushnir et al., 2015), and incorporate evidence quality when predicting belief revision (Confer et al., 2024; Magid et al., 2018) by demonstrating that children also expect belief updating to be based on prior belief strength. Together, these findings suggest that the assumption of this component of epistemic rationality is an integral part of our social cognition from early in life.

In Study 1, children judged changing strongly held beliefs to be more difficult than changing weakly-held beliefs. Importantly, however, children did not treat difficulty as equivalent to impossibility. Even four-year-olds (the youngest age group in our sample) judged that both strong and weak beliefs could, in principle, be revised. This distinction between difficulty and impossibility is noteworthy given that young children readily understand true epistemic constraints in other contexts: for example, that one cannot draw something without having seen it (Kushnir et al., 2015; Luo & Baillargeon, 2007) or decide to believe something that contradicts perceptual evidence (Confer et al., 2024). Our findings extend this work by showing that children also understand *graded* epistemic constraints: strong priors do not make belief change impossible, but they make it more difficult. The fact that four-year-olds in our study did not treat strong prior beliefs as creating impossibility offer suggestive evidence that children's representation of belief is probabilistic, rather than categorical or deterministic: they understand that the same belief can be held more or less strongly, and that the strength of a belief shapes the likelihood of its changing.

Study 2 extended these findings by examining children's predictions about how characters would respond to counterevidence. Beginning around age five, children predicted that a character with a strong prior was less likely to change their belief than one with a weak prior, even when the same evidence (i.e., testimony from a unanimous peer group) was presented. This finding further reveals that children see strong priors as exerting a constraining influence on belief change.

Study 2 also revealed emerging differences in how children reason about belief changes across domains. With age, children became less likely to predict belief change for moral norms in response to peer testimony. One potential explanation is that children evaluate not only the strength of prior beliefs, but also the appropriateness or relevance of different types of evidence for belief revision in different domains. Testimony from a consensus of peers may be seen as reliable evidence for empirical beliefs, but less persuasive for moral beliefs. In particular, moral beliefs may require justification from authoritative sources (Laupa & Turiel, 1986) and may be less amenable to testimonial transmission more broadly and more grounded in first-hand deliberation and reasoning (Hopkins, 2007; Li & Koenig, 2025). An alternative (though not necessarily incompatible) explanation is that children may reason differently about the malleability of certain kinds of beliefs based on their perceived connection to personal identity. Moral beliefs might, for instance, be judged as more central to people's identity or essence (Heiphetz, 2020; Heiphetz, Strohminger, & Young, 2017), thereby making them less susceptible to change through one-off testimonial evidence. A third consideration is that as children get older, they increasingly recognize individual variation in how people think (Carpendale & Chandler, 1996; Eisbach, 2004). This could include an expectation that different people respond to evidence differently–perhaps especially when it comes to moral or preferential beliefs. Though these domain-specific results are preliminary, they highlight the need for future research to explore not only how children reason about prior belief strength, but also how they evaluate the strength, quality, and domain-relevance of different types of evidence (see Fedyk et al., 2019).

A crucial question raised by our findings is why four-year-olds did not reliably distinguish between strong and weak prior beliefs. Several possibilities might explain this developmental pattern. One possibility is that sensitivity to prior strength emerges alongside broader shifts in children's understanding of belief. Some researchers have argued that before age five or six, children's success on false belief tasks may rely on simpler "perceptual access" reasoning (tracking what people have and have not seen) rather than genuine representation of beliefs as mental states with variable properties (Fabricius et al., 2021; Friedman & Petrashek, 2009). If children do not develop a full-fledged intuitive theory of belief until at least age five, this would explain why four-year-olds do not show sensitivity to prior strength in our tasks. Relatedly, children's ability to reason about graded epistemic notions may emerge around age five; prior work has found, for instance, that it is not until age four that children reliably distinguish between different degrees of inaccuracy (e.g., getting one versus three out of four questions wrong; Pasquini et al., 2007). A similar pattern appears in how young children reason about ignorance: Chen et al. (2015) found that young children equate ignorance with error rather than representing it as genuine uncertainty about what is true. Aboody et al. (2025) extended this finding by showing that it is not until age six that children consider that an ignorant agent could arrive at a correct answer by chance, suggesting that younger children do not yet represent the probabilistic space between knowing and not knowing. Similarly, Huang et al. (2019) found that children's ability to use the degree of certainty of another agent's knowledge to predict their subsequent behavior does not emerge stably until age six. Together, these findings point to a developmental shift from categorical to graded representations of others' epistemic states during the period between ages five and seven. Our findings may reflect this same shift: understanding that two characters who hold the same belief but differ in their degree of confidence will respond differently to counterevidence requires the kind of graded epistemic reasoning that these studies suggest is still emerging at age four to five.

A second possibility is methodological: our tasks required tracking

two characters with differing belief strengths and making comparative judgments, which may have challenged younger children's working memory. Additionally, our manipulations were language-heavy, relying on children understanding expressions like "believes it a lot" versus "believes it a little." However, we think task demands are unlikely to fully explain our results. Four-year-olds succeeded at basic components of the task: they judged that beliefs could change at above-chance rates in Study 1 and passed comprehension checks about belief strength. Moreover, the same developmental pattern emerged across two studies with different dependent measures and across multiple belief domains, suggesting the pattern reflects conceptual rather than purely task-related factors. Nevertheless, future research could use simplified paradigms such as single characters paradigms or nonverbal measures to help rule out task demands as a contributing factor.

A third possibility involves changes in how children represent the relationship between mental states and their consequences. Our task (particularly in Study 2) required children to understand that two characters who encounter the same evidence may nonetheless respond differently because their prior beliefs differed in strength. This capacity is related to what the literature on "interpretive" theory of mind has identified as a post–false-belief milestone: the understanding that the mind is actively constructive, such that identical informational input can lead to different beliefs depending on what the knower brings to the encounter (Carpendale & Chandler, 1996; Lalonde & Chandler, 2002). In related work, Eisbach (2004) found that the awareness that people will have different "trains of thought" in response to the same object because their individual backgrounds and knowledge shape what they think emerges between ages five and seven. Our finding that sensitivity to prior strength begins to emerge around age five may thus reflect the early stages of this broader developmental shift toward an interpretive theory of mind, even before the full capacity is firmly in place.

4.1. Limitations, connections, and future directions

Across both studies, we find evidence that by age five, children are sensitive to prior belief strength and use it to inform their judgments about the possibility, difficulty, and likelihood of belief change. An important future direction is to manipulate both prior strength and evidence strength within the same design to better understand when and how children integrate these components into their reasoning about belief revision. While the present studies focused on prior belief strength, children's open-ended justifications in Study 1 and domain-specific predictions in Study 2 also suggest sensitivity to the strength and relevance of counterevidence. Testing children's ability to integrate both factors will allow for a more complete picture of how their intuitive models of belief change align with rational updating principles. Future work could vary evidence strength (e.g., one peer vs. five peers) crossed with prior strength to examine how children integrate these components.

It is also important to note that Study 2 used testimonial counterevidence (i.e., a unanimous peer consensus) and our results therefore speak most directly to how children expect others to update in response to this type of second-hand, socially mediated evidence. Within a Bayesian framework, prior belief strength exerts its constraining influence primarily when evidence is probabilistic. However, testimonial evidence may be weaker than direct observation for some descriptive beliefs (though see Finiasz et al., 2024 and Macris & Sobel, 2017), and its perceived weight may vary by domain. This points to another important direction for future research: examining how the *type* of evidence influences children's predictions about belief revision. Future work could test this by comparing children's predictions when beliefs are challenged by testimonial versus observational counterevidence. It is possible that the influence of prior belief strength may be stronger for indirect evidence, and that children might expect even strongly held beliefs to yield in the face of direct, unambiguous evidence. Relatedly, future research should examine how the source of testimonial evidence influences children's predictions. Our studies used peer consensus because it could be applied uniformly across belief domains, and because it is often a powerful source of evidence for young children. However, peers may not be viewed as equally authoritative across all domains. Future work could compare children's predictions when beliefs are challenged by peer testimony versus testimony from adult authorities. If children's resistance to predicting moral belief revision reflects reasoning about who has authority over moral matters, they might respond differently to counterevidence from parents or teachers than from peers. This could help clarify whether domain differences in our findings reflect context-sensitive evaluation of evidential weight or more general intuitions about the stability of different belief types.

Another relevant finding regards children's base rate expectations about the likelihood of belief change. In Study 2, we found that even in the weak prior condition, children only predicted a character would change their belief around half the time when presented with counterevidence. Because our study was among the first to look at children's predictions about others' belief change, we did not have strong initial predictions about what the base rates of predicted change would be. In the only other study examining this question, Magid et al. (2018) found that children predicted belief change 63–77% of the time, depending on whether the evidence consisted of three or five randomly sampled observations. These findings raise important questions for future research about how much children expect others to change their beliefs in general, and what shapes those expectations. Some of the variation across studies likely reflects features of the situation such as evidence type (testimonial vs. observational), belief domain, and prior strength. But children's expectations may also be shaped by broader features of their developmental context. For example, children's experiences with parents' mental-state talk (Ruffman et al., 2002) and the degree to which adults in their lives model open belief revision (Mills et al., 2025) may inform what children come to see as typical or expected. Individual differences in emerging metacognitive and epistemic capacities such as intellectual humility (Danovitch et al., 2019; Porter et al., 2022) may further shape children's predictions about whether and when others will update their beliefs.

Our findings raise important questions about how children conceptualize belief itself. Though belief has traditionally been conceived of as a truth-oriented mental state that can take any kind of propositional content as its object (Schwitzgebel, 2006), a recent debate has challenged whether "belief" is a unified cognitive kind—both in terms of the attitude itself (Van Leeuwen, 2026), and in terms of people's lay theories about the attitude (Vesga et al., 2025; Westra, 2023). Pluralist views (e.g., Vesga et al., 2025; Westra, 2023) propose that some belief-attitudes are (understood as) primarily epistemic—formed to track truth—while others may be more "groupish," functioning to signal social affiliation, express identity, or align with group norms. From this perspective, beliefs may differ in how responsive they are to evidence, especially in morally or socially grounded contexts. Future work could explore when and whether young children begin to distinguish between epistemic versus non-epistemic beliefs, and whether there are "groupish" contexts in which prior belief strength and evidence strength are no longer the best predictors of how children reason about belief revision.

Finally, there are many remaining questions about how children's understanding of prior beliefs as epistemic constraints extends to other aspects of social cognition, particularly in their social learning and epistemic and moral evaluations. One direction for future research concerns whether children use their understanding of epistemic constraints to navigate disagreement. For instance, do children adjust their communicative strategies when they understand that a disagreeing interlocutor holds a strong versus a weak belief? Are they more likely to attempt persuasion when they perceive someone as holding a weaker belief? If so, then the inferences about belief revision documented here may serve as a foundation for children's developing capacities to communicate strategically and resolve disagreements (Köksal et al., 2025). Another direction concerns children's normative evaluations of

others' belief revision. Research suggests that children and adults readily make character judgments based on how others respond to evidence: individuals who refuse to update beliefs despite compelling counterevidence may be seen as stubborn, closed-minded, or intellectually arrogant, while those who are appropriately responsive to evidence may be seen as open-minded or intellectually humble (Bowes et al., 2026; Hagá & Olson, 2017; Porter et al., 2022). If children have expectations about rational belief revision, they may similarly evaluate others based on whether they conform to these expectations. For example, do children judge someone more negatively when they fail to update a weakly held belief in the face of strong evidence? Conversely, when someone does update a strongly held belief, do children see this as particularly praiseworthy given its difficulty? Prior work suggests that children tend to admire and learn from agents who overcome challenges or persist through difficulty (Leonard et al., 2017; Leonard et al., 2020), raising the question of whether they similarly value the epistemic virtue of revising strongly held beliefs in the face of compelling evidence, or whether they show forgiveness toward those who held problematic beliefs but demonstrated the capacity to change. These questions reveal that the role of prior belief strength is not important just for understanding theory of mind on its own–but also that it may have significant consequences for the development of social cognition more broadly.

4.2. Conclusions

Across two studies, we find that by around age five, children use information about prior belief strength to guide their reasoning about the possibility and difficulty of belief change, as well as the likelihood that others will revise their beliefs in light of counterevidence. We also found that children understand beliefs as mental states that are characteristically responsive to evidence. Importantly, these patterns emerged across diverse belief domains, suggesting that children's sensitivity to prior strength reflects a domain-general feature of their understanding of belief change. The emergence of this capacity around age five situates it within a broader developmental shift in representations of others' epistemic states. Taken together, these findings suggest that young children possess an understanding of belief revision that aligns with core principles of rational belief updating. Notably, children's reasoning reflects both conceptual flexibility and a surprisingly realistic view of belief revision: they recognize that beliefs can, in principle, be updated, yet they also understand that strongly held beliefs make belief change more difficult in practice.

Author note

The data and study materials that support the findings of the study is openly available in Open Science Framework and can be accessed at the following link: https://osf.io/ng5dr/

Studies 1 and 2 were preregistered, available at the following anonymized links: Study 1 (AsPredicted #133945): https://aspredicted.org/8cw8-c6bp.pdf; Study 2 (AsPredicted #176514): https://aspredicted.org/tshr-ycxp.pdf

CRediT authorship contribution statement

Pearl Han Li: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Laura K. Soter:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Data curation, Conceptualization. **Tamar Kushnir:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition.

Funding

This work was supported by the John Templeton Foundation (Grant ID 63351) and the Charles Lafitte Foundation Program in Psychological Research at Duke University.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Lanay Fuller, Sabreen Syed, Eleanor Myers and Reima Bash for their help with recruitment and data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2026.106668.

Data availability

The data and study materials are openly available in Open Science Framework (currently anonymized): https://osf.io/ng5dr/overview?view_only=cfdd18cbdc2c495184e53b9fed3ea677

References

Aboody, R., Lu, J., Denison, S., & Jara-Ettinger, J. (2025). Six-year-olds, but not younger children, consider the probability of being right by chance when inferring others' knowledge. *Child Development, 96*(5), 1765–1776.

Atance, C. M., & Caza, J. S. (2018). "Will I know more in the future than I know now?" preschoolers' judgments about changes in general knowledge. *Developmental Psychology, 54*(5), 857.

Baer, C., Engelmann, J. M., & Kidd, C. (2025). Children use the relative confidence of people with conflicting perspectives to form their own beliefs. *Developmental Science, 28*(4), Article e70027.

Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*(4), 946–964. https://doi.org/10.2307/1131035

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *Fitting linear mixed-effects models using lme4 [R]*.

Bowes, S. M., Novick, K., Lourenco, S. F., & Tasimi, A. (2026). Do children value intellectual humility over intellectual arrogance? *Developmental Psychology, 62*(3), 611.

Browne, C. A., & Woolley, J. D. (2004). Preschoolers' magical explanations for violations of physical, social, and mental laws. *Journal of Cognition and Development, 5*(2), 239–260. https://doi.org/10.1207/s15327647jcd0502_4

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*(4), 1686–1706.

Chen, Y., Su, Y., & Wang, Y. (2015). Young children use the "ignorance= getting it wrong" rule when predicting behavior. *Cognitive Development, 35*, 79–91.

Chernyak, N., Kang, C., & Kushnir, T. (2019). The cultural roots of free will beliefs: How Singaporean and U.S. children judge and explain possibilities for action in interpersonal contexts. *Developmental Psychology, 55*(4), 866–876. https://doi.org/10.1037/dev0000670

Chernyak, N., Kushnir, T., Sullivan, K. M., & Wang, Q. (2013). A comparison of American and Nepalese children's concepts of freedom of choice and social constraint. *Cognitive Science, 37*(7), 1343–1355. https://doi.org/10.1111/cogs.12046

Confer, J. A., Schleihauf, H., & Engelmann, J. M. (2024). Children and adults' intuitions of what people can believe. *Child Development, 95*(2), 447–461. https://doi.org/10.1111/cdev.13988

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological Science, 20*(3), 372–377. https://doi.org/10.1111/j.1467-9280.2009.02291.x

Corriveau, K. H., & Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. *Developmental Psychology, 46*(2), 437.

Danovitch, J. H., Fisher, M., Schroder, H., Hambrick, D. Z., & Moser, J. (2019). Intelligence and neurophysiological markers of error monitoring relate to children's intellectual humility. *Child Development, 90*(3), 924–939.

Dennett, D. C. (1988). Précis of the intentional stance. *Behavioral and Brain Sciences, 11*(3), 495–505. https://doi.org/10.1017/S0140525X00058611

Eason, A. E., Doctor, D., Chang, E., Kushnir, T., & Sommerville, J. A. (2018). The choice is yours: Infants' expectations about an agent's future behavior based on taking and receiving actions. *Developmental Psychology, 54*(5), 829.

Eisbach, A. O. D. (2004). Children's developing awareness of diversity in people's trains of thought. *Child Development, 75*(6), 1694–1707.

Fabricius, W. V., Gonzales, C. R., Pesch, A., Weimer, A. A., Pugliese, J., Carroll, K., & Spinrad, T. L. (2021). Perceptual access reasoning (PAR) in developing a

representational theory of mind. *Monographs of the Society for Research in Child Development, 86*(3), 7–154.
Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.
Fedyk, M., Kushnir, T., & Xu, F. (2019). Intuitive epistemology: Children's theory of evidence. In *Advances in experimental philosophy of science*. Bloomsbury Academic. https://doi.org/10.5040/9781350068896.0010.
Feldman, R., & Conee, E. (1985). Evidentialism. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 48*(1), 15–34.
Finiasz, Z., Gelman, S. A., & Kushnir, T. (2024). Testimony and observation of statistical evidence interact in adults' and children's category-based induction. *Cognition, 244*, Article 105707.
Flavell, J. H., & Green, F. L. (1999). Development of intuitions about the controllability of different mental states. *Cognitive Development, 14*(1), 133–146. https://doi.org/10.1016/S0885-2014(99)80021-5
Flavell, J. H., Green, F. L., & Flavell, E. R. (1998). The mind has a mind of its own: Developing knowledge about mental uncontrollability. *Cognitive Development, 13*(1), 127–138. https://doi.org/10.1016/S0885-2014(98)90024-7
Friedman, O., & Petrashek, A. R. (2009). Children do not follow the rule "ignorance means getting it wrong". *Journal of Experimental Child Psychology, 102*(1), 114–121.
Gärdenfors, P., & Makinson, D. (1988). Revisions of knowledge systems using epistemic entrenchment. Proceedings of the. In *2nd conference on theoretical aspects of reasoning about knowledge* (pp. 83–95).
Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology, 60*(1), 115–140.
Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in Cognitive Sciences, 7*(7), 287–292.
Gerstenberg, T., & Tenenbaum, J. B. (2017). Intuitive theories. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning, Oxford library of psychology* (2017 online ed.). Oxford Academic https://doi-org.proxy.lib.umich.edu/10.1093/oxfordhb/9780199399550.013.28.
Goodman, N. D., Baker, C., Bonawitz, E., Mansinghka, V. K., Gopnik, A., Wellman, H. M., … Tenenbaum, J. B. (2006). Intuitive theories of mind: A rational approach to false belief. *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society, 6*, 1382–1387.
Gopnik, A., & Astington, J. (1988). Children's understanding of representational change and its relation to the under-standing of false belief and the appearance–reality distinction. *Child Development, 59*, 26–37.
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
Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2011). Bayes and blickets: Effects of knowledge on causal induction in children and adults. *Cognitive Science, 35*(8), 1407–1455. https://doi.org/10.1111/j.1551-6709.2011.01203.x
Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107*(20), 9066–9071.
Hagá, S., & Olson, K. R. (2017). 'If I only had a little humility, I would be perfect': Children's and adults' perceptions of intellectually arrogant, humble, and diffident people. *The Journal of Positive Psychology, 12*(1), 87–98.
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.
Haun, D. B., & Tomasello, M. (2011). Conformity to peer pressure in preschool children. *Child Development, 82*, 1759–1767.
Heiphetz, L. (2020). The development and consequences of moral essentialism. In , *Vol. 59. Advances in child development and behavior* (pp. 165–194). Elsevier. https://doi.org/10.1016/bs.acdb.2020.05.006.
Heiphetz, L., Gelman, S. A., & Young, L. L. (2017). The perceived stability and biological basis of religious beliefs, factual beliefs, and opinions. *Journal of Experimental Child Psychology, 156*, 82–98. https://doi.org/10.1016/j.jecp.2016.11.015
Heiphetz, L., Spelke, E. S., Harris, P. L., & Banaji, M. R. (2013). The development of reasoning about beliefs: Fact, preference, and ideology. *Journal of Experimental Social Psychology, 49*(3), 559–565. https://doi.org/10.1016/j.jesp.2012.09.005
Heiphetz, L., Strohminger, N., & Young, L. L. (2017). The role of moral beliefs, memories, and preferences in representations of identity. *Cognitive Science, 41*(3), 744–767. https://doi.org/10.1111/cogs.12354
Ho, M. K., Saxe, R., & Cushman, F. (2022). Planning with theory of mind. *Trends in Cognitive Sciences, 26*(11), 959–971. https://doi.org/10.1016/j.tics.2022.08.003
Hopkins, R. (2007). What is wrong with moral testimony? *Philosophy and Phenomenological Research, 74*(3), 611–634. https://doi.org/10.1111/j.1933-1592.2007.00042.x
Huang, Z., Hu, Q., & Shao, Y. (2019). Understanding others' knowledge certainty from inference and information-seeking behaviors in children. *Developmental Psychology, 55*(7), 1372.
Jara-Ettinger, J., Floyd, S., Tenenbaum, J. B., & Schulz, L. E. (2017). Children understand that agents maximize expected utilities. *Journal of Experimental Psychology: General, 146*(11), 1574–1585. https://doi.org/10.1037/xge0000345
Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition, 140*, 14–23. https://doi.org/10.1016/j.cognition.2015.03.006
Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science, 17*(9), 757–758. https://doi.org/10.1111/j.1467-9280.2006.01778.x
Jern, A., Chang, K. K., & Kemp, C. (2014). Belief polarization is not always irrational. *Psychological Review, 121*(2), 206–224. https://doi.org/10.1037/a0035941
Kimura, K., & Gopnik, A. (2019). Rational higher-order belief revision in young children. *Child Development, 90*(1), 91–97. https://doi.org/10.1111/cdev.13143
Koenig, M. A., & Harris, P. L. (2005a). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*(6), 1261–1277. https://doi.org/10.1111/j.1467-8624.2005.00849.x
Koenig, M. A., & Harris, P. L. (2005b). The role of social cognition in early trust. *Trends in Cognitive Sciences, 9*(10), 457–459. https://doi.org/10.1016/j.tics.2005.08.006
Köksal, Ö., Saffran, A., & Sodian, B. (2025). Let me show why you are wrong": The origins of scientific argumentation, its development, and cognitive predictors. *Developmental Psychology, 62*(3), 513.
Kushnir, T., & Gopnik, A. (2005). Young children infer causal strength from probabilities and interventions. *Psychological Science, 16*(9), 678–683.
Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43*(1), 186–196. https://doi.org/10.1037/0012-1649.43.1.186
Kushnir, T., Gopnik, A., Chernyak, N., Seiver, E., & Wellman, H. M. (2015). Developing intuitions about free will between ages four and six. *Cognition, 138*, 79–101. https://doi.org/10.1016/j.cognition.2015.01.003
Kushnir, T., & Koenig, M. A. (2017). What I don't know won't hurt you: The relation between professed ignorance and later knowledge claims. *Developmental Psychology, 53*(5), 826–835. https://doi.org/10.1037/dev0000294
Kushnir, T., Wellman, H. M., & Gelman, S. A. (2008). The role of preschoolers' social understanding in evaluating the informativeness of causal interventions. *Cognition, 107*(3), 1084–1092.
Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21*(8), 1134–1140. https://doi.org/10.1177/0956797610376652
Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology, 20*(2–3), 163–198.
Langenhoff, A. F., Engelmann, J. M., & Srinivasan, M. (2023). Children's developing ability to adjust their beliefs reasonably in light of disagreement. *Child Development, 94*(1), 44–59. https://doi.org/10.1111/cdev.13838
Laupa, M., & Turiel, E. (1986). Children's conceptions of adult and peer authority. *Child Development, 57*(2), 405–412. https://doi.org/10.2307/1130596
Leonard, J. A., Garcia, A., & Schulz, L. E. (2020). How adults' actions, outcomes, and testimony affect preschoolers' persistence. *Child Development, 91*(4), 1254–1271.
Leonard, J. A., Lee, Y., & Schulz, L. E. (2017). Infants make more attempts to achieve a goal when they see adults persist. *Science, 357*(6357), 1290–1294.
Li, P. H., & Koenig, M. A. (2025). Appealing to consequences, or authority? The influence of explanations on children's moral judgments across two cultures. *Cognition, 254*, Article 105994. https://doi.org/10.1016/j.cognition.2024.105994
Lin, H. (2024). Bayesian epistemology. In E. N. Zalta, & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy (summer 2024)*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2024/entries/epistemology-bayesian/.
Liu, R., & Xu, F. (2022). Can adults revise their core beliefs about objects? *Proceedings of the Annual Meeting of the Cognitive Science Society, 44*(44). https://escholarship.org/uc/item/3d12384w.
Liu, R., & Xu, F. (2023). Can adults revise their core beliefs about agents? *Proceedings of the Annual Meeting of the Cognitive Science Society, 45*(45). https://escholarship.org/uc/item/7966283s.
Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science, 358*(6366), 1038–1041.
Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition, 105*(3), 489–512.
Macris, D. M., & Sobel, D. M. (2017). The role of evidence diversity and explanation in 4- and 5-year-olds' resolution of counterevidence. *Journal of Cognition and Development, 18*(3), 358–374. https://doi.org/10.1080/15248372.2017.1323755
Magid, R. W., Yan, P., Siegel, M. H., Tenenbaum, J. B., & Schulz, L. E. (2018). Changing minds: Children's inferences about third party belief revision. *Developmental Science, 21*(2), Article e12553.
Mahr, J. B., & Csibra, G. (2021). A short history of theories of intuitive theories. In *A life in cognition: Studies in cognitive science in honor of Csaba Pléh* (pp. 219–232). Cham: Springer International Publishing.
Masnick, A. M., Klahr, D., & Knowles, E. R. (2017). Data-driven belief revision in children and adults. *Journal of Cognition and Development, 18*(1), 87–109. https://doi.org/10.1080/15248372.2016.1168824
McLoughlin, N., Finiasz, Z., Sobel, D. M., & Corriveau, K. H. (2021). Children's developing capacity to calibrate the verbal testimony of others with observed evidence when inferring causal relations. *Journal of Experimental Child Psychology, 210*, Article 105183. https://doi.org/10.1016/j.jecp.2021.105183
Mills, C. M., Danovitch, J. H., & Quintero, N. B. (2025). Does intellectual humility transmit intergenerationally? Examining relations between parent and child measures. *Journal of Experimental Psychology: General., 154*(8), 2272.
Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43*(5), 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216
Pesch, A., Ridge, K. E., Suárez, S., McMyler, B., & Koenig, M. A. (2022). Evaluations of epistemic and practical reasons for belief in a predominantly White U.S. sample of preschoolers. *Journal of Experimental Child Psychology, 223*, Article 105499. https://doi.org/10.1016/j.jecp.2022.105499
Porter, T., Elnakouri, A., Meyers, E. A., Shibayama, T., Jayawickreme, E., & Grossmann, I. (2022). Predictors and consequences of intellectual humility. *Nature Reviews Psychology, 1*(9), 524–536.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology, 1*(4), 223–235. https://doi.org/10.1038/s44159-022-00037-z

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development, 73*(3), 734–751.

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*(4), 1054–1070. https://doi.org/10.1111/1467-8624.00334

Schleihauf, H., Herrmann, E., Fischer, J., & Engelmann, J. M. (2022). How children revise their beliefs in light of reasons. *Child Development, 93*(4), 1072–1089. https://doi.org/10.1111/cdev.13758

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology, 40*(2), 162–176. https://doi.org/10.1037/0012-1649.40.2.162

Schwitzgebel, E. (2006). Belief. In *Stanford encyclopedia of philosophy*. https://plato.stanford.edu/entries/belief/?TB_iframe=true&width=370.8&height=658.8#WhatItBeli.

Sebastián-Enesco, C., Guerrero, S., & Enesco, I. (2020). What makes children defy their peers? Chinese and Spanish preschoolers' decisions to trust (or not) peer consensus. *Social Development, 29*(2), 494–508. https://doi.org/10.1111/sode.12416

Shtulman, A., & Phillips, J. (2018). Differentiating "could" from "should": Developmental changes in modal cognition. *Journal of Experimental Child Psychology, 165*, 161–182. https://doi.org/10.1016/j.jecp.2017.05.012

Sobel, D. M. (2004). Exploring the coherence of young children's explanatory abilities: Evidence from generating counterfactuals. *British Journal of Developmental Psychology, 22*(1), 37–58. https://doi.org/10.1348/026151004772901104

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review, 120*(4), 779–797. https://doi.org/10.1037/a0034191

Sobel, D. M., & Letourneau, S. M. (2015). Children's developing understanding of what and how they learn. *Journal of Experimental Child Psychology, 132*, 221–229.

Sobel, D. M., Li, J., & Corriveau, K. H. (2007). "They danced around in my head and I learned them": Children's developing conceptions of learning. *Journal of Cognition and Development, 8*(3), 345–369.

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28*(3), 303–333. https://doi.org/10.1207/s15516709cog2803_1

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition, 95*(1), 1–30.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*(7), 309–318. https://doi.org/10.1016/j.tics.2006.05.009

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

Titelbaum, M. G. (2022). *Fundamentals of Bayesian epistemology 1: Introducing Credences*. Oxford University Press.

Van Leeuwen, N. (2026). The trinity and the light switch: Two faces of belief. E. Schwitzgebel & J. Jong (eds.), The nature of belief. Oxford University Press, (Forthcoming).

Vesga, A., Van Leeuwen, N., & Lombrozo, T. (2025). Evidence for multiple kinds of belief in theory of mind. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001765

Wainryb, C., Shaw, L. A., Langley, M., Cottam, K., & Lewis, R. (2004). Children's thinking about diversity of belief in the early school years: Judgments of relativism, tolerance, and disagreeing persons. *Child Development, 75*(3), 687–703.

Wang, Z. (2015). Theory of mind and children's understanding of teaching and learning during early childhood. *Cogent Education, 2*(1), Article 1011973.

Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology, 15*(6), 728–755. https://doi.org/10.1080/17405629.2018.1435413

Wellman, H. M., Kushnir, T., Xu, F., & Brink, K. A. (2016). Infants use statistical sampling to understand the psychological world. *Infancy, 21*(5), 668–676. https://doi.org/10.1111/infa.12131

Westra, E. (2023). Symbolic belief in social cognition. *Philosophical Perspectives, 37*(1), 388–408. https://doi.org/10.1111/phpe.12196

Williams, B. (1973). Deciding to believe. In B. Williams (Ed.), *Problems of the self* (pp. 136–151). Cambridge University Press.

Xu, F., & Kushnir, T. (2013). Infants are rational constructivist learners. *Current Directions in Psychological Science, 22*(1), 28–32.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review, 114*(2), 245–272. https://doi.org/10.1037/0033-295X.114.2.245

Young, A. G., Alibali, M. W., & Kalish, C. W. (2012). Disagreement and causal learning: Others' hypotheses affect children's evaluations of evidence. *Developmental Psychology, 48*(5), 1242.