Young children strategically adapt to unreliable social partners

Katherine Adams Shannon, Aneesa Conine-Nakano, Willem E Frankenhuis, Michael C Frank, Hyowon Gweon

Child Development, 2026, 796–810
https://doi.org/10.1093/chidev/aacag012
Published: March 24, 2026
Empirical Article

Society for Research in Child Development OXFORD

Young children strategically adapt to unreliable social partners

Katherine Adams Shannon[1], Aneesa Conine-Nakano[2], Willem E. Frankenhuis[3,4], Michael C. Frank[1], and Hyowon Gweon[1,*]

[1]Department of Psychology, Stanford University, Stanford, CA, United States
[2]Department of Psychology, University of Chicago, Chicago, IL, United States
[3]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, Netherlands
[4]Max Planck Institute for the Study of Crime, Security and Law, Freiburg im Breisgau, Germany
*Corresponding author: Department of Psychology, Stanford University, 450 Jane Stanford Way, Building 420, Stanford, CA 94305, United States. Email: hyo@stanford.edu

Abstract

The effectiveness of social learning depends on whether learners receive help when they need it. In four preregistered studies, U.S. 4–6-year-olds (*N* = 244; 54% female, 48% Asian, 2% Black, 9% Hispanic/Latino, 24% Multiracial/Other, 27% White) interacted with an adult who either did or did not follow through on promised help. Experiment 1 tested the effect of reliable vs. unreliable help on children's future task choice; Experiment 2 examined its effect on children's help-seeking and exploration of a novel toy. Children's learning goals and strategies were modulated by the past reliability of help, suggesting that seemingly maladaptive decisions—such as avoiding a hard task—may be adaptive responses that balance the reliability of help against the utility of exploring alone.

Keywords social learning, adaptation, cognitive development

Lay summary

Receiving help when needed is crucial for learning, but it is not a given. The current study asked whether 4- to 6-year-old children strategically modulate their learning depending on the reliability of their social partners. Children's learning choices and strategies depended on whether an adult kept promises to help. When help was reliable, children were more likely to tackle harder tasks and ask for assistance; when help was unreliable, they avoided difficult tasks and explored independently. These patterns suggest that even seemingly "maladaptive" behaviors—like avoiding a challenge—can be reconstrued as adaptive responses. Children can weigh the likelihood of receiving help against what they can achieve on their own and make adaptive learning decisions depending on their social environment.

Early learning is both active and social; children learn not only by independently exploring their environment but also by learning from others. Children can flexibly switch between these strategies, asking for help selectively when a learning task is too complex to solve alone (Cluver et al., 2013; Gweon & Schulz, 2011). Yet, receiving help when requested is not a given: Teachers might promise help but fail to deliver, or parents may be unavailable to help or even unresponsive. How do children navigate learning in environments where help is reliable or unreliable?

All children experience variation in the reliability of help across time, contexts, and social partners. For example, there is substantial heterogeneity in caregiving at all socioeconomic levels: between families living in the same community (Kuchirko & Tamis-LeMonda, 2019) and even among children within a single household (von Stumm & Latham, 2018). Outside of their family environment, young children also need and request help from their social partners in a range of contexts, especially in both formal and informal learning contexts. Thus, regardless of their socioeconomic environment, children may experience considerable variation in the reliability of help. Such variability may have direct downstream consequences for their learning decisions and strategies. However, this causal relationship is challenging to test in real-world contexts, because it is difficult both ethically and practically to intervene on such a fundamental feature of children's environment.

In the current paper, we take an alternative approach by developing a small-scale simulation of reliable or unreliable help within

Handling Editor: Dave Sobel
Editor: Shauna Cooper
Received: June 24, 2025. **Revised:** November 3, 2025. **Accepted:** December 19, 2025

an experimental procedure (e.g., see Kidd, Palmeri, & Aslin, 2013). This method provides a model system for exploring how past social experience can influence future learning through children's own adaptive decision-making (Frankenhuis & de Weerth, 2013; Suor et al., 2017). By studying the direct impact of the past reliability of help on subsequent learning in a lab setting, we may gain useful insights that can inform our understanding of how the reliability of children's social environments affects early learning in real-world contexts.

Young children's sensitivity to reliability

The motivation for experimentally studying the relationship between reliability of help and learning in a lab setting comes from a body of research in cognitive development on early social learning. Prior work suggests that young children are selective social learners; rather than blindly trusting others, they prefer to learn from reliable informants (see Harris et al., 2018; Mills, 2013; Sobel & Kushnir, 2013). In this literature, the key to reliability is a belief about the informant's trustworthiness in terms of their knowledge and informativeness (i.e., epistemic reliability): A reliable informant can be trusted to provide information that is beneficial for learning. Research suggests that children's representation of epistemic reliability is not only informed by individual traits such as age (Jaswal & Neely 2006) or markers of social groups such as accent (Corriveau et al. 2013; Kinzler et al., 2011), but also by the informants' past informativeness, such as their accuracy or completeness of demonstration in prior interactions (Birch et al., 2008; Gweon & Asaba, 2018; Koenig & Harris, 2005).

Critically, this literature has shown that children are also *adaptive* social learners, who flexibly change their strategies depending on others' past epistemic reliability. This notion of adaptivity is often grounded in theories of learning that assume that children's adaptive responses approximate a *rational* or even *optimal* response to the statistical structure of the environment (Anderson, 1990; Tenenbaum et al., 2011; Xu, 2019). For instance, preschoolers can track subtle differences in relative accuracy between teachers (Pasquini et al., 2007), and discount prior testimony if new evidence shows that a previously informative partner has become uninformative (Corriveau & Harris, 2009 ; Ronfard et al., 2017; Scofield & Behrend, 2008). Children also adapt their own behaviors when their social partner has provided insufficient information in the past. For instance, when an adult leaves ambiguity in their referential expression, children spontaneously seek additional information by looking at them (Hembacher et al., 2020). Furthermore, after seeing that a teacher was under-informative in the past (e.g., demonstrated just 1 of 4 functions of a toy), children engage in compensatory exploration when the teacher demonstrates a new toy (Gweon et al., 2014).

Taken together, this literature has shown that young children readily track, and adapt to, others' reliability as informants. However, although prior work on pedagogical reasoning has highlighted both knowledgeability and helpfulness as two key properties of an effective informant (e.g., Bonawitz et al., 2011; Gweon, 2021; Gweon et al., 2014), the notion of helpfulness in this literature primarily concerns the informant's *intention* to help—their pedagogical goal to provide useful information, in particular—rather than the *availability* of help. Given recent work suggesting that children's beliefs about how much a teacher knows (i.e., epistemic trust) may be distinct from their beliefs about whether a teacher offers help or keeps their promises (i.e., interpersonal trust; Pesch & Koenig, 2023), children's sensitivity to epistemic reliability may not necessarily generalize to other forms of reliability in their social environment.

Some prior work has examined how certain forms of environmental or interpersonal reliability affect children's subsequent behaviors. One study, for instance, used a variant of the classic delay-of-gratification "marshmallow" task (Kidd et al., 2013), showing that children's tendency to wait for a promised reward was influenced by the experimenter's prior history of delivering on their promise. Before the marshmallow task, the experimenter either showed herself to be reliable or unreliable based on whether or not she gave the child a promised reward. Children in the unreliable condition had much shorter wait times than those in the reliable condition, choosing the one guaranteed marshmallow over the promise of two marshmallows later. In a similar paradigm where the experimenter acted untrustworthy toward another adult—lying about having destroyed their art—children also showed shorter wait times (Michaelson & Munakata, 2016). While short wait times in the classic marshmallow task are taken as an index for poor self-control (Mischel et al., 1988; Mischel & Ebbesen, 1970), children's behavior in these unreliable and untrustworthy conditions reflects an adaptive response to their environment: When a promised reward is unlikely to be delivered, attaining an immediate but smaller reward is a reasonable low-risk, low-reward strategy. These studies and other replications, which show similar if attenuated effects, demonstrate that even a brief experimental manipulation can induce children to infer that their current social environment is unreliable, and adapt their behavior accordingly (Lee & Carlson, 2015; Moffett et al., 2020).

The current study

In the current work, our use of the term "reliability" in the context of helping is informed by both literatures—prior work on selective social learning and epistemic reliability (e.g., Harris et al., 2018; Mills, 2013; Sobel & Kushnir, 2013), as well as studies on children's responses to interpersonal or environmental reliability (e.g., Kidd et al., 2013; Michaelson & Munakata, 2016)—focusing on the idea that a reliable social partner can be trusted to respond to requests for help and act in ways that benefit the recipient. Studying children's sensitivity and adaptation to this form of reliability is particularly important for at least two reasons. First, most prior research on social learning—including studies that manipulate epistemic reliability—assumed learning environments where adults or informants are generally responsive. Second, prior work on environmental reliability has focused on manipulating reliability in terms of promise-keeping or general trustworthiness, with downstream effects that are largely limited to self-regulatory behaviors. Given the importance of adults' help in scaffolding early learning, manipulating the reliability of help may lead to rational adaptations in how children learn: in particular, how children select future tasks and navigate the balance between self-guided exploration and help-seeking. Consider, for instance, a young child whose parent promised to help on a challenging task, but then is unresponsive to their request for help. If the parent is routinely unavailable or breaks their promise to help, the child might learn something about their environment: Help may not be available when they need it. As a result, they may discount help-seeking as a useful learning strategy in the future.

Such responses may influence at least two kinds of decisions that are critical for learning. First, a child in an unreliable environment may be more reluctant to try challenging tasks than a child in a reliable environment. Avoiding challenging tasks—which might result in missed learning opportunities—has often been characterized as unmotivated or even maladaptive (De Castella et al., 2013; Dicintio & Gee, 1999; Stipek, 1996). However, in an unreliable social environment where bids for help often go unanswered, it actually might be *better* to avoid a challenging task, especially if it cannot be achieved without assistance, and instead choose an easier task that can be achieved alone. Thus, avoiding challenging tasks, which might appear suboptimal at face value, may be contextually appropriate when a child's past experience is taken into account (Frankenhuis & Del Giudice, 2012). By considering the learner's social context, we can characterize this sort of behavior as a rational adaptation to environmental constraints, rather than a deficit of the learner (Ellis et al., 2023). There is, however, a trade-off: Forgoing challenging opportunities in unreliable contexts, though reasonable, can have downstream consequences that limit learning and growth, especially if the child's learning affordances improve over time (e.g., when entering a new school in which teachers do provide reliable help).

Second, a child in an unreliable environment may also use different learning strategies than a child in a reliable environment. Imagine, for instance, that seeking help from an unresponsive caregiver, who routinely fails to help when needed, often results in no information gained. Then, the child might rely more on self-guided exploration rather than seeking help or instruction, even when exploration may not lead to meaningful discoveries. Such exploration can appear maladaptive when the task is especially difficult or impossible to complete alone, since it may result in wasted effort or frustration rather than progress. Yet it is also a reasonable response that can still offer some opportunities for learning when help is unavailable. In contexts where independence is rewarding (e.g., a task that children can figure out on their own with sufficient exploration), this strategy could, in fact, turn out to be even more beneficial than receiving help (Bonawitz et al., 2011). Consistent with this idea, recent work shows that children's independent exploration can foster persistence and support deeper learning (Leonard et al., 2021; Liquin & Gopnik, 2022). However, it remains unknown how children weigh the likelihood of receiving help against the perceived utility of exploration when choosing how to learn.

On these theoretical, empirical, and applied grounds, our goal is to take an experimental approach to examine whether children adjust their reliance on social learning based on the past reliability of help. To this end, we randomly assigned children to interact with a reliable or unreliable social partner (the experimenter). This approach allows us to test the causal relationship between children's prior social experiences—whether their bids for help were answered—and their subsequent learning behaviors. Our central hypothesis is that children infer the reliability of help based on past experience, then use this social reasoning to make decisions about what and how to learn next. From this, we predict that children who interact with an unreliable experimenter should, in subsequent interactions, (1) choose less challenging tasks, and (2) persist more in their exploration (instead of asking for help), compared to children who interact with a reliable experimenter.

We tested these two predictions through four preregistered experiments with 4- to 6-year-old children. We focus on ages 4 to 6 because children in this age range can track and adapt their learning behaviors based on the informativeness of their social partners (see Gweon, 2021, for a review), and are going through an important transition into school where they navigate learning in new social environments (Blair, 2002; Cook & Coley, 2017). In Experiments 1A and 1B, we ask whether children consider the past reliability of help in deciding which goal to pursue next: whether to choose a more challenging but more rewarding task or to settle for an easier, less rewarding task. In Experiments 2A and 2B, we ask whether children consider the past reliability of help in deciding their learning strategies: whether to seek help or explore alone to learn how a novel puzzle toy works. In the Supporting Information (SI), we also report a separate study that yielded null effects and informed the design of Experiments 2A and 2B. All study procedures were approved by Stanford University's Institutional Review Board. All experiments were preregistered; data and analysis code for all experiments are available on OSF (https://osf.io/a8sk5/).

Experiment 1A

Experiment 1 tested our first hypothesis: whether children adapt their future goals based on the past reliability of help. In both Experiments 1A and 1B, we followed a sequential sampling procedure, described below, in which the sample size was not fixed in advance but determined by repeatedly evaluating the key hypothesized effect as data collection progressed, until either a predetermined evidentiary criterion or final N was met (Etikan, 2016). Given the relatively long duration of each experimental session and hence the challenges of collecting larger samples, this allowed us to stop early in the presence of sufficient evidence for or against the hypothesis, or continue to collect a full sample in the absence of this evidence.

Method

Participants

The sequential sampling procedure (see Results) yielded a final $N = 60$ (4;1–6;10 years, $M_{age} = 5.50$ years, 50% female, 37% Asian, 4% Black, 8% Hispanic/Latino, 14% Multiracial, 6% Other, 39% White). Data were collected between March and December 2022. As in all subsequent experiments, participants were recruited from a local children's museum in the US. Of the 53 families who reported highest parent education attainment, 81% completed graduate degrees, 4% completed some graduate training, and 15% completed college. Participants were typically developing, and heard English at least 50% of the time at home per caregiver report. Fifteen children were excluded due to not completing the experiment ($n = 7$; e.g., child wanting to stop early), experimenter error ($n = 7$), or parent-reported developmental delay ($n = 1$). Caregivers provided written consent and children verbal assent for participation. Sampling, exclusion, and analytic procedures were preregistered at https://aspredicted.org/xr5g-776s.pdf.

Materials and procedure

This between-subjects experimental design had two phases (see Figure 1). In the history phase, children were presented with a project to decorate a piece of paper to display in a cup and take home. The experimenter showed the child a single brown or gray crayon in a jar, explaining that normally this is the only crayon available to use. The experimenter pretended to have difficulty opening the

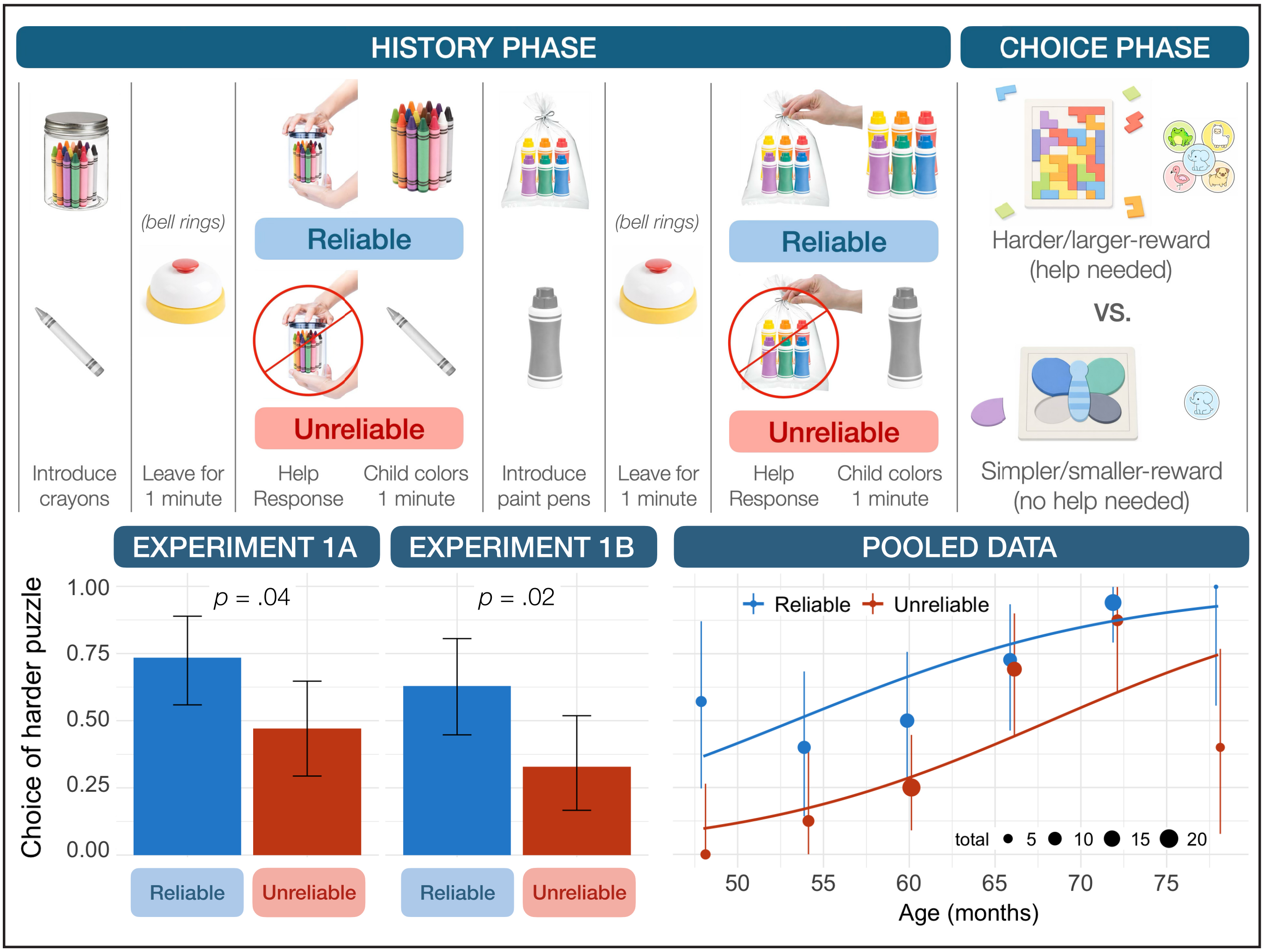

Figure 1 Schematic of the procedure (top) and results (bottom) in Experiment 1. Top: The history phase consisted of two episodes of reliable or unreliable help, followed by the choice phase where children selected a puzzle to complete. The puzzles were visible in Experiment 1A and obscured in Experiment 1B. Bottom: Proportion of children choosing the harder puzzle (left) and choice of harder puzzle by age in the pooled sample (right). Dots represent means for a given age group in 6-month bins; dot sizes represent the sample size (n) per bin; trend lines are best-fitting logistic regression curves. Error bars in both plots are 95% bootstrapped CIs. Exact *p* values in logistic regression are shown; see main text for details.

jar but succeeded, exclaiming it was a bit stuck. This sequence established the experimenter's competence at opening stuck jars. The experimenter then revealed another jar of cool, colorful crayons, explaining that a friend in the room next door let them borrow these to use instead, and put the single brown or gray crayon away. She then explained that she had to do some work nearby while the child colored, but the child could ring a bell on the table for any help if needed. This explicit promise to help made it clear to the child that assistance was available in this new context and demonstrated how to request it. She handed the jar of colorful crayons to the child and exited the room.

The child soon discovered that the jar of colorful crayons was also stuck closed (secretly glued shut), prompting them to ring the bell for help. One minute after leaving the room, the experimenter returned. In the Reliable condition, the experimenter opened the jar (by secretly switching the glued jar for an identical unglued one) and helped the child access the colorful crayons. In the Unreliable condition, the experimenter explained that she was too busy to help while the child rang the bell and now the friend next door asked for the jar of colorful crayons back, offering the single brown or gray crayon instead to the child. In both conditions, the child worked on coloring for one minute. Thus, children in both conditions tried a difficult task they could not achieve themselves (i.e., opening a stuck jar) but only those in the Reliable condition received help and were able to use colorful crayons; children in the Unreliable condition received no such help, and as a consequence, had to use the single crayon.

After coloring with crayons, a similar process was repeated with another set of art supplies: a bag of colorful paint pens that was secretly glued shut. As with the crayons, in the Reliable condition, the experimenter returned and opened the bag, whereas in the Unreliable condition, the experimenter said they were too busy to help and the bag needed to be returned, offering a single brown or gray paint pen instead. Thus, by the end of the history phase, children had two consecutive interactions in which they asked for help, and the experimenter either responded with promised help or did not respond with promised help. The "too busy" explanation provided a plausible reason for not responding to bids for help while also implying that the experimenter's priority was their own work and that the same decision might reoccur, rather than a temporary interruption (e.g., taking a phone call). When the experimenters did not help, they used an apologetic tone and maintained a positive, encouraging demeanor throughout both conditions.

The choice phase followed the history phase. The same experimenter told the child that the next activity was playing with a puzzle. The experimenter pointed to the simpler/smaller-reward puzzle and said, "This puzzle is really easy. Most kids can do it by themselves. If you finish this puzzle, you'll get one sticker." She then pointed to the harder/larger-reward puzzle and said, "This puzzle is really hard. Most kids need help to do it. If you finish this puzzle, you'll get five stickers." Finally, the experimenter prompted the child to choose a puzzle to complete and promised help again: "Now listen carefully, I have to go do some more work in the room next door. But I'm really good at puzzles. So, you can call me if you need any help by pressing this bell, okay? So, you can only choose one puzzle: Which puzzle would you like to do?" The left vs. right presentation position of the puzzles was counterbalanced across conditions.

After the child made the selection, the experimenter "realized" that the session was almost out of time and gave the child the simpler puzzle, explaining that she would rather stay and do this puzzle with the child than work more. All participants finished the simpler puzzle with help and encouragement from the experimenter and received five stickers upon its completion. This procedure allowed us to reestablish reliability for participants in the Unreliable condition.

Results

Our key dependent variable of interest was children's binary choice of either the simpler/smaller-reward puzzle or harder/larger-reward puzzle. Our hypothesis (H_1) was a negative effect of the Unreliable condition on choice of harder/larger-reward puzzle. We conducted a logistic regression analysis in R of the form `choice ~ condition`.

We computed a one-sided Bayes factor (BF) analysis using the `BFpack` package (Mulder et al., 2021) with a contingency table BF, assessing the data being more likely under the hypothesis of a condition effect (H_1) relative to the null hypothesis (H_0: no difference between conditions; assuming Cauchy priors). While we hypothesized finding strong evidence of a condition effect ($BF_{10} > 10$), following guidelines for reporting Bayesian analyses (van Doorn et al., 2021) also support interpreting $BF_{10} > 5$ as moderate evidence and $BF_{10} > 3$ as weak evidence. This analysis informed our sequential sampling procedure. We tested an initial sample of 10 children per condition, and then evaluated the BF for a condition effect after each day of testing. Stopping criteria was set at a $BF_{10} > 10$ in favor of H_1, a $BF_{01} > 3$ in favor of H_0, or at a maximum sample size of $n = 30$ per condition ($N = 60$ total).

Results show that children in the Reliable condition had about three times higher odds of choosing the harder/larger-reward puzzle ($n = 22$ of 30) than children in the Unreliable condition ($n = 16$ of 30; $OR = 3.14$; 95% CI [1.09, 9.64]; see Figure 1). Although the regression p value was significant ($p = .04$), our primary evaluation of the condition effect produced a $BF_{10} = 2.18$, suggesting that the data were only about twice as likely under the hypothesis of a condition effect than under the null hypothesis of no condition difference. This result did not reach the threshold ($BF_{10} > 3$) for weak but meaningful evidence of a condition effect (note that such divergence between frequentist and Bayesian approaches is common; see Benjamin et al. (2018) for a review). We did not reach our sampling stopping criterion of $BF_{10} > 10$, thus our sample size reflects the maximum $N = 60$.

As a planned secondary analysis, we conducted a follow-up regression including a main effect of age (in months, centered) and an interaction of age and condition. In the interaction model, Unreliable condition ($\beta = -1.12$, 95% CI [−2.36, 0.06], $p = .07$), age ($\beta = 0.06$, 95% CI [−0.03, 0.17], $p = .20$), and the interaction term ($\beta = 0.12$, 95% CI [−0.05, 0.34], $p = .20$) were not significant, and the condition effect had a $BF_{10} = 1.36$. In a non-preregistered analysis, we dropped the interaction term to increase statistical power and clarify the main effects, given that the interaction was non-significant and not central to our hypothesis. Both Unreliable condition ($\beta = -1.27$, 95% CI [−2.51, −0.13], $p = .03$) and older age ($\beta = 0.11$, 95% CI [0.03, 0.20], $p = .01$) were significant predictors of puzzle choice; condition effect had a $BF_{10} = 2.37$.

Discussion

Experiment 1A yielded suggestive evidence that children's prior experience with a more or less reliable social partner, as well as their age, may be related to their choice of future tasks. However, these results did not meet our minimum Bayes factor threshold.

Why might the negative effect of unreliable help on the choice of the harder/larger-reward puzzle be smaller than hypothesized? One key feature of Experiment 1A is that the learning goals—hard and easy puzzles—were visible to the participant. Although the experimenter described the puzzles as easy to do alone or hard and generally requiring help, children could also see the puzzles and make their own judgments of whether or not they would need the experimenter's help to complete each puzzle. In particular, as reflected in children's tendency to choose harder puzzles with age, older children may have judged the presented puzzles to be generally easier to complete alone than younger children. Furthermore, because the puzzles were commercially available, some children may also have played with similar puzzles in the past and already knew how to solve them. We designed Experiment 1B to address these concerns.

Experiment 1B

Experiment 1B was identical to Experiment 1A, aside from how the puzzles were presented. Instead of showing the actual puzzles, the simpler/smaller-reward and harder/larger-reward puzzles were concealed in separate boxes, thus eliminating any visual cues about the puzzle design or difficulty. Additionally, because the experimenters in these experiments were not blind to condition, we tried to minimize any unintended bias at the time of the final choice by having a coder without knowledge of the condition try to predict the child's choice based solely on watching video of the experimenter's face, hands, and voice during the choice phase; the child's voice was audible, but the child was not visible. Cohen's κ value of .19 (95% CI [−.061, .45]) suggested that the experimenter's behavior did not reliably predict the child's choice.

Methods

Participants

The same sequential sampling procedure was implemented, leading to our final sample size of $N = 60$ (4;0–6;10 years, $M_{age} = 5.36$

years, 53% female, 52% Asian, 2% Black, 4% Hispanic/Latino, 15% Multiracial, 30% White). Data collection occurred between August 2022 and January 2023. Of the 50 families who reported highest parent education attainment, 78% completed graduate degrees, 2% completed some graduate training, 16% completed college, 2% completed some college, and 2% completed high school. Participants were typically developing. Per caregiver report, all participants heard English at least 25% of the time; 94% of participants heard English at least 50% of the time. All participants belonged to separate households. Eight children were excluded due to the experiment stopping early ($n=3$) and experimenter error ($n=5$). Sampling, exclusion, and analytic procedures were preregistered at https://aspredicted.org/86n5-tx6q.pdf.

Materials and procedure

Materials and procedure were identical to Experiment 1A, with the exception that the puzzles were hidden in unmarked white boxes. The boxes differed in size, providing an extra clue to help children tell the boxes apart as the experimenter described the puzzles inside as either hard (larger box) or easy (smaller box).

Results

Except where noted, all sampling and analytic procedures were the same as in Experiment 1A. Note that, due to an unequal distribution of ages across conditions, we continued data collection despite reaching our sequential sampling stopping criterion of $BF_{10}=19.17$ at $n=28$. Age was better distributed across conditions for the subsequent BF analysis at $n=34$, and the criterion of $BF_{10}>10$ was no longer met. Data collection continued for a sample size of $N=60$.

In Experiment 1B with puzzles obscured (see Figure 1), children in the Reliable condition had about three and a half times higher odds of choosing the harder/larger-reward puzzle ($n=19$ of 30) than children in the Unreliable condition ($n=10$ of 30; OR = 3.45; 95% CI [1.22, 10.35], $p=.02$). This analysis produced a $BF_{10}=3.50$ in favor of H_1, which we interpret as weak evidence of a condition effect (van Doorn et al., 2021) although it indicates less support than our hypothesis of a strong condition effect ($BF_{10}>10$). In our planned secondary analysis investigating the effects of age on puzzle choice, we find no significant interaction of age and condition ($\beta=-0.05$, 95% CI [−0.24, 0.12], $p=.54$). Unreliable condition ($\beta=-1.51$, 95% CI [−2.86, −0.29], $p=.02$) and age ($\beta=0.17$, 95% CI [0.06, 0.33], $p=.01$) were both significant predictors of puzzle choice, and the condition effect had a $BF_{10}=3.86$. In a non-preregistered analysis dropping the interaction term, both Unreliable condition ($p=.02$) and age ($p<.001$) remained significant main effects, and the condition effect had a $BF_{10}=4.35$.

Pooled results

In a post hoc, non-preregistered analysis, we pooled participants across Experiments 1A and 1B to increase power to detect an age-by-condition interaction, adding experiment as a fixed effect (see Figure 1, bottom right). In the pooled data ($N=120$), we again find no significant interaction of age and condition ($\beta=0.03$, 95% CI [−0.09, 0.15], $p=.61$), nor experiment ($\beta=-0.48$, 95% CI [−1.35, 0.35], $p=.26$). Unreliable condition ($\beta=-1.39$, 95% CI [−2.28, −0.55], $p=.002$) and age ($\beta=0.11$, 95% CI [0.04, 0.20], $p=.004$) were both significant predictors of puzzle choice. The condition effect had a $BF_{10}=26.30$, indicating strong evidence for a condition effect and an expected increase given the doubled sample size. Both Unreliable condition ($p=.002$) and age ($p<.001$) remained significant when the interaction term was removed, and the condition effect had $BF_{10}=28.18$.

Discussion

Although Experiment 1B was largely identical to Experiment 1A, the main difference was that the puzzles were not visible at the time of choice. As in Experiment 1A, children in the Unreliable condition were more likely to choose the smaller, easier puzzle than those in the Reliable condition ($p=.02$). The Bayes factor in Experiment 1B ($BF_{10}=3.50$) met our preregistered threshold and thus provides modest evidence for a condition effect. In both experiments, we also observed an increasing preference for the more challenging puzzle with age. Finally, a post hoc pooled analysis showed strong evidence for the condition and age effects across Experiments 1A and 1B. Taken together, these results support our first hypothesis: Children who experienced unreliable help were more likely to choose a less challenging task than those who experienced reliable help. Our findings are consistent with the possibility that children's choice of future learning goals is adaptive with respect to the reliability of their environment: Children in the Unreliable condition might have chosen an easier task despite the smaller reward because they believe they are less likely to complete the harder, more-rewarding puzzle without help.

One question that emerges from these findings is: What happens when children who experienced unreliable help must face a challenging task? In subsequent experiments (Experiments 2A and 2B), we test our second hypothesis: Children might adapt their help-seeking and exploration strategies based on the past reliability of help. Specifically, those who previously experienced reliable help may be quicker to seek assistance, whereas those who experienced unreliable help may prefer to explore rather than asking for help. We test this hypothesis using an exploratory play paradigm across two different task contexts that also varied in the utility of exploration, asking how children weigh the likelihood of receiving help against the perceived effectiveness of their own actions.

Experiment 2A

In Experiment 2A, we tested our hypothesis using a multi-sided novel toy with many possible functions. The task was to discover how to activate lights on top of the toy. This design was informed by a separate preregistered study with a simple four-button toy, which yielded no condition difference, likely because the toy was not interesting enough to motivate sustained exploration (see SI, Appendix A). In that earlier study, the help bell was placed on the table next to the toy, allowing children to seek help while continuing to explore rather than choosing between strategies. These limitations motivated the present version with a multifunction toy and the bell placed away from the toy. Similar to that earlier study, the toy was inert and could not be activated by the child; we alter this decision in Experiment 2B.

Methods

Participants

We used the same sequential sampling procedure as Experiment 1, leading to a final $N=32$ (4;1–6;9 years, $M_{age}=5.33$ years, 56%

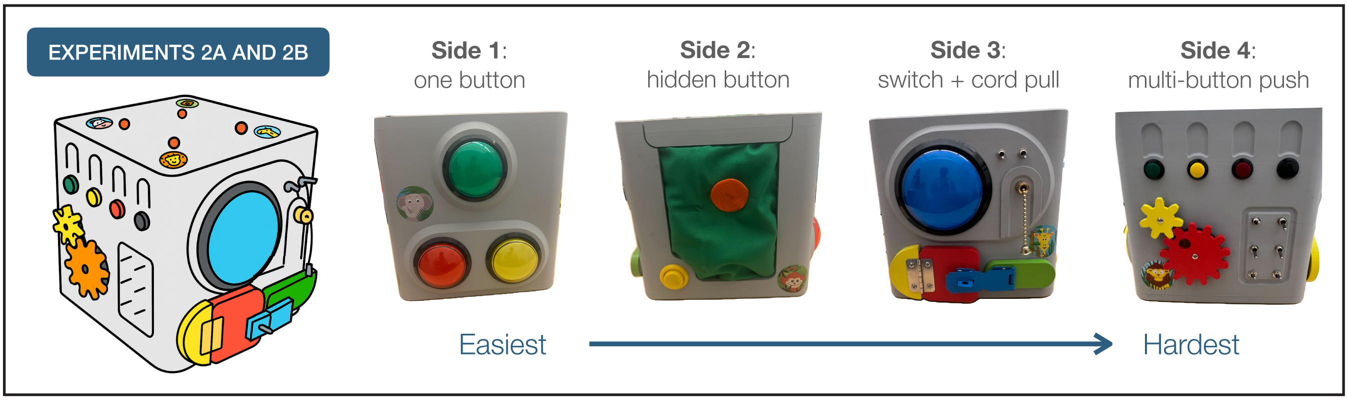

Figure 2 Schematic (left) and images (right) of the toy used in Experiments 2A and 2B. In Experiment 2A, the toy was inert and did not activate during free play. In Experiment 2B, the toy was functional; each of its four sides featured a puzzle that varied in difficulty.

female, 68% Asian, 15% Hispanic/Latino, 8% Multiracial, 24% White). Data collection occurred between March and May 2024. Of the 27 families who reported highest parent education attainment, 70% completed graduate degrees, 4% completed some graduate training, 22% completed a 4-year college, and 4% completed high school. Participants were typically developing, and heard English at least 50% of the time at home per caregiver report. Nine children were excluded due to the experiment stopping early ($n = 7$) or equipment error ($n = 2$; e.g., video did not record). Caregivers provided written consent and children verbal assent for participation. Sampling, exclusion, and analytic procedures were preregistered at https://aspredicted.org/zspw-qrmt.pdf.

Materials and procedure

The history phase which establishes the reliability of the experimenter was the same as in Experiment 1 (see Figure 1). The free play phase followed the history phase. The same experimenter told the child that the next activity was playing with a novel puzzle toy.

The experimenter presented to the child a multifunction box with four sides, each containing many different affordances (see Figure 2). However, all functions of the toy were disabled (i.e., no switch, button, or pull chain caused any effect). The experimenter explained, "This special box has four sides with puzzles that make the lights on top turn on. If you can figure it out, you get a bunch of cool stickers! How does that sound?" Finally, the experimenter explained that they needed to leave the room and promised help: "Now listen carefully, I have to finish my work in the room next door. If you need any help, you can call me by pressing the bell and I'll give you some hints, okay? Alright, I'm going to go do my work and you can call me if you need any help." The experimenter then left the room for 3 minutes. The bell (the signal for the experimenter to help) was located on a bench across the room so that children had to disengage from exploring to make a bid for help. The experimenter did not answer any bids for help during this play session, so all children completed the 3-minute free play phase alone without further interaction. Importantly, unbeknownst to the children, the toy was inert and therefore impossible to activate.

After the child completed the 3-minute session, the experimenter returned and inquired if the child had solved the puzzle (none had by design as the toy was nonfunctional). The experimenter "realized" that the toy must be broken and offered another simpler puzzle toy to the child which they solved together to play music. Thus, all children discovered the causal mechanism needed to operate a toy with help and encouragement from the experimenter, and received five stickers upon its completion. This closing procedure allowed us to reestablish the experimenter's reliability in both conditions.

Video annotation

Video data were coded using Datavyu (www.datavyu.org), an open-source video coding software that allows for frame-by-frame navigation and precise time-stamping of event onsets and offsets. Coders marked the timing and presence of several key behaviors during the 3-minute free play phase in which the child was alone in the room, which began when the experimenter closed the door to leave and reopened the door to return. These behaviors included: (1) *whether the child bid for help* (yes or no); (2) *duration until first bid for help* (if no bid for help occurred, the absence was noted and the duration was imputed as the maximum 3 minutes of the play session); and (3) *episodes of exploration*, defined as periods when the child was actively engaging with the toy, from the moment they began using their hands to explore until they stopped touching the toy for more than 5 seconds. If the child started a new behavior after taking their hands off the toy, this ended the exploration episode even if the child returned to touching in less than 5 seconds, and started a new episode. The duration of the exploration episodes was summed into a total duration of exploration.

All videos were double-coded by two trained coders. Coders could flag timestamps for review by an expert coder who resolved any ambiguities. Original and reliability coding files for each participant were examined for error or discrepancies (e.g., coder marked the onset and offset of an episode but indicated the behavior as absent rather than present). Inter-rater reliability was assessed on all double-coded videos prior to discrepancy resolution.

Coders demonstrated perfect rank-order agreement for the duration until first bid for help (Spearman's $\rho = 1.00$) with a mean absolute difference (MAD) of 1.52 seconds, and 93.8% of observations within the 4-second threshold. For cumulative exploration time, reliability was also high ($\rho = 0.98$, MAD $= 3.67$ seconds), with 90.6% of values within the 8-second threshold. All discrepancies exceeding these thresholds were resolved by an expert coder, who provided the final data for analysis.

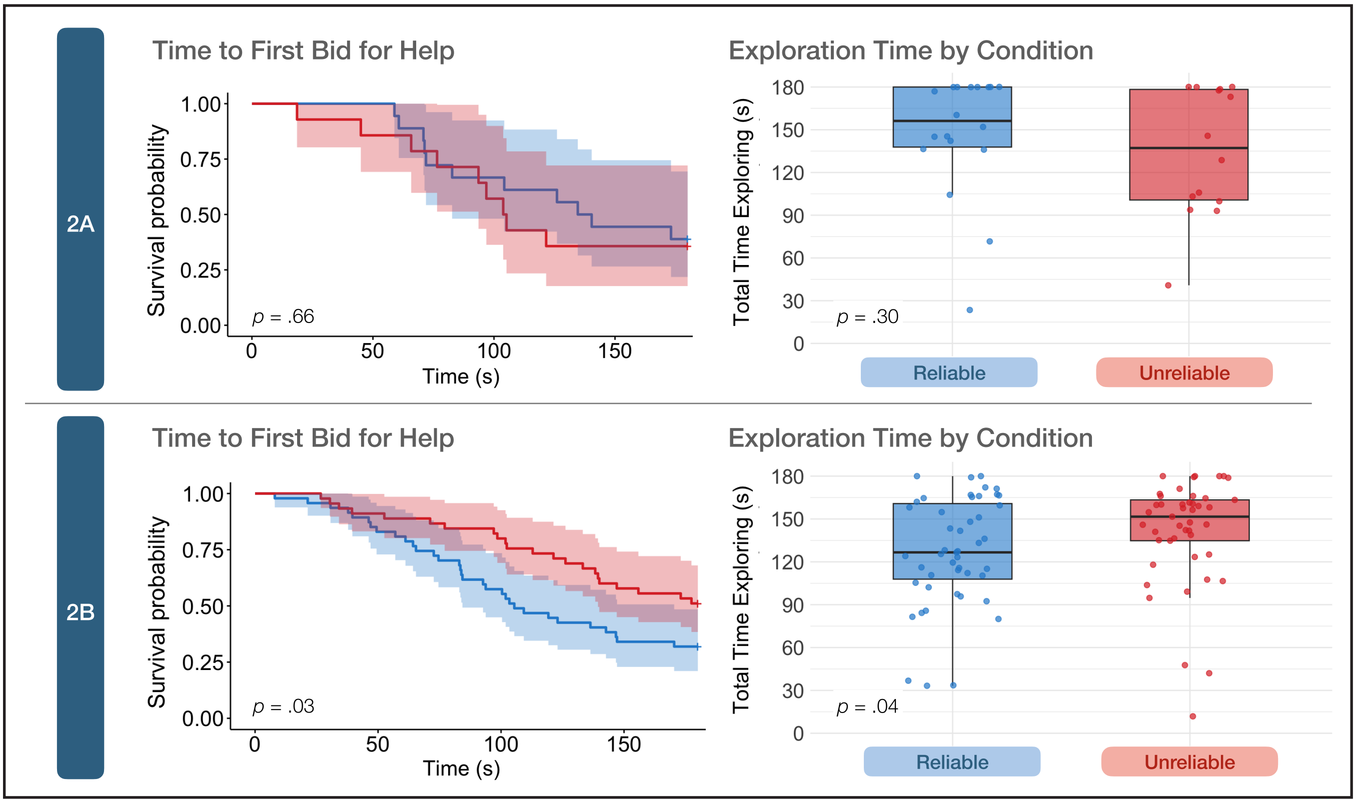

Figure 3 Results from Experiment 2A (top) and Experiment 2B (bottom). The plots show survival curves for each condition, representing the proportion of children who did not yet make a bid for help at each time point during free play (left) and the amount of time children spent exploring the toy in each condition (right). Exact *p* values are shown (see main text for details).

Results

Our key preregistered dependent variables of interest were the presence or absence of a first bid for help and, if such a bid occurred, the duration in seconds until it did. These dependent variables were analyzed jointly using survival analysis, which accounts for the potential of heavily right-censored data (i.e., children may never make a bid for help and thus would not have a "duration until first bid for help" data point). The analysis was conducted using the `survival` package in R (Therneau, 2023) of the form `survival(time_to_first_help, help_sought) ~ condition`. In addition to reporting *p* values from the Cox proportional hazards regression, we used the `baymedr` package (Linde & van Ravenzwaaij, 2023) to compute a one-sided Bayes factor assessing evidence for the hypothesis that the Unreliable condition results in a lower probability of requesting help and an overall shorter time to request help (H_1). We defined the default (truncated Gaussian) prior on effects with a mean of zero and a standard deviation of 1. As in Experiment 1, we interpreted a $BF_{10} > 10$ as strong evidence, $BF_{10} > 5$ as moderate evidence, and $BF_{10} > 3$ as weak evidence for a condition effect.

The proportion of children with at least one bid was 61% (11 of 18) in the Reliable condition and 64% (9 of 14) in the Unreliable condition (see Figure 3). Among those who bid for help, the mean time to first bid was 99.52 s (SD = 38.55) and 80.77 s (SD = 32.78) in the Reliable and Unreliable conditions, respectively. The Cox proportional hazards regression showed no significant effect of condition on the likelihood of bidding for help, $\beta = 0.20$, $SE = 0.45$, $z = 0.44$, $p = .66$, $HR = 1.22$, 95% CI [0.50, 2.95]. The Bayes factor comparing the alternative hypothesis of a negative effect of Unreliable condition on help-seeking (H_1) to the null (H_0) was $BF_{10} = 0.32$; in other words, the data were about 3 times more likely under H_0 than under H_1 ($BF_{01} = 3.16$). This result met our sampling stopping criterion of $BF_{01} > 3$ in favor of the null hypothesis at a sample size of $N = 32$.

In a non-preregistered analysis, we also compared the total duration of toy exploration across conditions during the 3-minute free play phase using a Wilcoxon rank sum test due to non-normality of the data. Children in the Unreliable condition ($M = 134.30$ s, $SD = 45.43$) did not differ significantly in exploration time compared to those in the Reliable condition ($M = 147.43$ s, $SD = 43.23$), $W = 154$, $p = .30$. Because the condition effect was not significant, we did not conduct follow-up analyses with covariates.

Discussion

In the current experiment, the toy was disabled so children never experienced any successful discovery across the 3 minutes of free play. The null result we observed suggests that, even though the multifunction toy offered many affordances to explore, the expected utility of exploration was too low to justify forgoing the opportunity to receive help. Thus, even children in the Unreliable condition, who may have been less certain that the experimenter would respond to their bids for help, may have resorted to asking for help nonetheless. This is plausible especially because the experimenter, in both conditions, offered children help before leaving the room.

One possibility is that if children *know* they can discover at least some functions (i.e., the expected utility of exploration is sufficiently high), they might be more likely to respond differently depending on the past reliability of help: those who experienced reliable help may be quick to make a bid for help, whereas those who doubt the experimenter would come may continue to rely on exploration. Experiment 2B addresses this possibility.

Experiment 2B

Experiment 2B tested whether children adapt their help-seeking and exploration strategies based on the past reliability of help. Rather than using a disabled toy, we reconfigured the multifunction toy from Experiment 2A so that each of the four puzzle sides was solvable. Based on a pilot study where children explored each of the four sides, we expected that all four sides could, in principle, be solved independently by the child, but their difficulty levels would range from very easy (i.e., nearly all children could discover; henceforth Side 1) to very challenging (i.e., almost no children would discover, henceforth Side 4; see Figure 2). We hypothesized that children in the Unreliable condition would be slower to, and less likely to, seek help, and that they would explore more and discover more causal mechanisms of the toy.

Methods

Participants

Experiment 2B was preregistered to have either a maximum sample size of $N=100$ or to have data collected until September 30, 2024, the conclusion of our study period, whichever came first. Data were collected between July and September 2024. Our final sample size was $N=92$ (4;1–6;11 years, $M_{age}=5.49$ years, 57% female, 46% Asian, 1% Black, 11% Hispanic/Latino, 31% Multiracial, 5% Other, 17% White). Of the 81 families who reported highest parent education attainment, 68% completed graduate degrees, 1% completed some graduate training, 21% completed a 4-year college degree, 5% completed a 2-year college degree, and 2% completed some college training. Participants were typically developing, and heard English at least 50% of the time at home per caregiver report. An additional 24 children were excluded due to not completing the experiment ($n=22$) and experimenter error ($n=2$). One additional child was excluded post hoc based on the consensus of the authors and three unaffiliated researchers that the child was scared of touching the toy. Caregivers provided written consent and children verbal assent for participation. Sampling, exclusion, and analytic procedures were preregistered at https://aspredicted.org/mdvf-33cj.pdf.

Materials and procedure

The history phase which establishes the reliability of the experimenter was the same as in Experiment 1 and Experiment 2A (see Figure 1). The free play phase that followed the history phase was nearly identical to Experiment 2A, including the use of the multifunction box. The only change implemented in Experiment 2B is that the box was turned on, such that children could discover the causal mechanisms to activate each of the puzzle sides through their own exploration (see Figure 2). The experimenter said, "This special box has four sides. Each of the sides has a puzzle that makes a light on top turn on and play a sound. Your goal is to solve as many puzzles as you can! For each puzzle you figure out, you get a bunch of cool stickers!" Children were presented the box with the easiest Side 1 facing them, although they were not given any information about the difficulty of the puzzles or further instructions on how to explore the box. After the 3-minute free play session, the experimenter helped the child activate any puzzle sides that were not solved and gave the child five stickers. This closing procedure allowed us to reestablish the experimenter's reliability in both conditions.

Video annotation

Experiment 2B used largely the same video coding, discrepancy resolution, and inter-rater reliability calculation procedures described in Experiment 2A. Videos were coded for (1) *whether the child bid for help*, (2) *duration until first bid for help*, and (3) *episodes of exploration*, except that the exploration episodes were coded for each side of the toy, and their duration was summed into a total duration of exploration. When children explored multiple sides simultaneously (e.g., touching two sides at once), overlapping episodes were collapsed so that time was not double-counted in the cumulative total. We also coded (4) *the total number of sides solved* where the child successfully discovered the causal mechanism that activated a given side (total possible sides solved is 4). In addition to the video annotation in Datavyu, an additional coder also reported for each participant (1) *the total number of causal functions the child interacted with* (out of 21); and (2) *whether the child attempted a complex combination* defined as interacting with two or more causal affordances at the same time.

Trained coders demonstrated perfect rank-order agreement for the duration until first bid for help (Spearman's $\rho=1.00$, MAD$=1.42$ seconds) with 95.7% of observations within a 4-second threshold. For the total duration of exploration, reliability was high ($\rho=0.95$, MAD$=7.73$ seconds), with 70.7% of observations within 8 seconds. Discrepancies exceeding these thresholds were reviewed and resolved by an expert coder.

Results

Except where noted, all sampling and analytic procedures were the same as in Experiment 2A. The proportion of children with at least one bid was 68% (32 of 47) in the Reliable condition and 49% (22 of 45) in the Unreliable condition (see Figure 3). Among those who bid for help, the mean time to first bid was 84.43 s ($SD=39.92$) and 104.35 s ($SD=46.35$) in the Reliable and Unreliable conditions, respectively. The Cox proportional hazards regression revealed a significant effect of condition on the likelihood of bidding for help, $\beta=-0.61$, $SE=0.28$, $z=-2.21$, $p=.03$. The hazard ratio for the Unreliable condition was 0.54, 95% CI [0.31, 0.93], indicating that children in this condition were significantly less likely and slower to bid for help compared to those in the Reliable condition. The Bayes factor comparing the alternative hypothesis of a negative effect of Unreliable condition on help-seeking (H_1) to the null (H_0) was $BF_{10}=5.33$, meaning the data were about five times more likely under H_1 than under the null.

We examined differences in exploration behavior across conditions during the 3-minute free play phase. A Wilcoxon rank-sum test revealed that children in the Unreliable condition spent significantly more time exploring the toy ($M=141.15$ s, $SD=37.02$) than those in the Reliable condition ($M=126.91$ s, $SD=37.92$), $W=798$, $p=.04$. Note, we deviated here from our preregistration because we had inappropriately preregistered the use of survival analysis to model cumulative exploration time. Survival analysis is appropriate for looking at time until a discrete event, not cumulative time.

Most children solved the easiest Side 1 (100% in Reliable, 93% in Unreliable), and success rates declined across subsequent, increasingly complex sides (Side 2: 45% and 53%; Side 3: 17% and 20%; Side 4: 4% and 4%, in Reliable and Unreliable conditions, respectively). On average, children in the Reliable condition solved 1.66 sides of the toy ($SD=0.73$), while those in the Unreliable condition solved 1.71 sides ($SD=0.87$). This difference was not significant (preregistered analysis; $\beta=0.05$, $SE=0.21$, $z=0.25$, $p=.80$, $OR=1.05$, 95% CI [0.69, 1.61]), suggesting that increased exploration in the Unreliable condition did not lead to more discovery.

A series of non-preregistered analyses further suggested that unreliable help prompted broader exploration of the functional novel toy, however. Children in the Unreliable condition ($M=17.73$, $SD=4.69$) explored a greater number of potential causal affordances (e.g., buttons, switches, cords, gears; out of 21 possible) than those in the Reliable condition ($M=15.60$, $SD=5.00$), $W=803.5$, $p=.04$. They were also significantly more likely to attempt complex combinations—defined as simultaneously testing two or more potential mechanisms (e.g., pushing two buttons at once)—with 62% (28 of 45) of children in the Unreliable condition and 38% (18 of 47) in the Reliable condition, $OR=2.65$, 95% CI [1.14, 6.16], $p=.02$. Follow-up analyses with covariates are reported in the Supporting Information (see SI, Appendix B).

Discussion

With a multifunction toy that allowed children to make meaningful discoveries, manipulating the reliability of help led to a number of differences in children's learning strategies. All but two participants (98%) solved at least one side—often within the first few seconds of the free play phase—providing clear evidence that the effectiveness of their own actions was high. Compared to the children in the Reliable condition, fewer children in the Unreliable condition asked for help, and if they did, they tended to do so later. Notably, children in the Unreliable condition also explored the toy longer, explored more broadly, and were more likely to attempt complex actions. While these behaviors did not lead to a difference in learning (i.e., number of sides solved), this may be due to the low variability in the number of remaining functions given that almost all children solved one side and about half solved two sides.

Taken together, in Experiment 2B, we found a clear effect of the past reliability of help on children's subsequent exploration and help-seeking. All of these differences in behavior are consistent with children adapting their strategy to maximize their expected utility: Facing a lower probability of receiving help, they were less likely to request help and instead relied more on self-guided exploration.

General discussion

The current project examined whether 4- to 6-year-olds adapted their learning decisions (e.g., try an easy or hard task; Exp. 1) and strategies (e.g., explore or ask for help; Exp. 2) based on the past reliability of receiving help. Experiments 1A and 1B together show that children who experienced unreliable help would forego a challenging but more rewarding task and choose an easier but less rewarding task instead, compared with those who received help when needed. Experiments 2A and 2B asked whether children would explore more and ask for help less when they had previously experienced unreliable help. Experiment 2A, which used a multifunction toy that children could not activate on their own, did not yield evidence for this hypothesis. However, when that same toy was turned on and thus functional, Experiment 2B found that those exposed to unreliable help explored longer and more broadly, and asked for help later and less, compared to children who experienced reliable help. Overall, despite some mixed evidence, these data provide support for our central hypothesis: Children can infer the reliability of help based on past experience and strategically adapt their learning decisions and strategies. Furthermore, these results are consistent with the possibility that children's adaptations rationally balance the likelihood of receiving help against the expected utility of trying a task without help: If help from social partners was unreliable in the past, they seek assistance less and do more on their own in the future, at least when they find that their own actions can lead to discovery.

Rational adaptation to unreliable partners

Why did children in Experiment 1 make different learning decisions depending on the reliability of help? In the Reliable condition, the harder/larger-reward puzzle was a reasonable choice because children could harness a responsive partner who could provide help if needed. Within an explore/exploit framework (Cohen et al., 2007; Frankenhuis & Gopnik, 2023), choosing the harder/larger-reward puzzle may be viewed as the child exploring their skills ("Can I solve this to earn five stickers?") with little risk of failure if the task was too difficult. In the Unreliable condition, however, that would have been a risky choice. Given that their bids for help were likely to go unanswered, children played it safe by attempting the simpler puzzle on their own to get only one sticker rather than risk gaining nothing by trying—and ultimately failing to solve—a very difficult task (i.e., choosing the simpler puzzle exploited the child's known skills: "I can solve this to earn one sticker."). In other words, what counts as a more adaptive choice differed depending on the reliability of help, and children made choices that reflected this pattern.

In a similar vein, children's exploration and help-seeking behaviors in Experiment 2B can also be explained as a reasonable response to both the expected utility of exploration and the likelihood of receiving help. While children who experienced unreliable help may forgo challenging tasks in the first place, once children are already in a challenging task context, the reasonable response is to rely on their own exploration rather than wait for help that may not come after all. Critically, however, if it is unlikely that exploration would yield any meaningful discoveries, children may resort to seeking help even when the prospect of receiving it is low. This is consistent with what we observed in Experiment 2A, where the toy was inactive and children experienced no indication that their exploration would lead to success.

Taken together, these studies recast learning decisions that may have been previously considered maladaptive as an adaptive response to the social environment (e.g., see Frankenhuis et al., 2020; Frankenhuis & Del Giudice, 2012). In the context of our experiment, these choices can be explained as an inference informed by utility-based social reasoning (see Baker et al., 2017; Jara-Ettinger et al., 2016, for utility-based reasoning as a computational framework for social cognition). Although children around this age can reason about another learner's expected utilities to decide what to teach (Bridgers et al., 2020), little work has directly investigated how utility-based social reasoning may account for children's ability to adapt their own learning depending on the predicted availability of support. The current work suggests that children consider the behavior of their social partners when reasoning about the costs and benefits of their own learning. Such reasoning may not necessarily involve explicit awareness of their own utilities, but nonetheless inform children's implicit decision-making processes.

These adaptations, even when well-reasoned, can necessarily contribute to differences in learning outcomes. This is in part because children may be miscalibrated in their assessment of the expected utility of exploration and the reliability of help in the future. For instance, in Experiments 1A and 1B, the two puzzles were explicitly marked as hard and easy, and it was implied that the hard puzzle may be too difficult for children to solve on their own. Yet, in real-world contexts, tasks are rarely marked in terms of whether children can do it on their own, and children may be miscalibrated in their estimates of task difficulty. By choosing a simpler task after experiencing unreliable help, children may have limited their opportunity for learning, especially when they could have benefitted from trying the harder task. Similarly, children's predictions about the future availability of help may not always be accurate. In particular, children who face variable reliability—either due to unpredictability within an environment or large discrepancies across different environments (e.g., home vs. school)—might miscalibrate their anticipation of help. In such cases, an adaptation that was once beneficial can become ineffective if the learner fails to update their expectations and behavior as their social environment changes (Frankenhuis et al., 2018; Frankenhuis & Del Giudice, 2012; Patterson et al., 1992). Children's behaviors in Experiment 2B highlight this point: Children who experienced unreliable help continued to explore without seeking help on puzzles that were too difficult for most children to solve. This behavior would be a reasonable adaptation in a truly unreliable environment, but in a context where help is readily available, such a response would have been a maladaptive one. At the same time, a child who is used to receiving help may lose valuable opportunities for self-guided discovery if they expect to be helped and forgo trying something on their own. Indeed, recent work suggests that 4- to 5-year-old children who interacted with an experimenter who took over (rather than taught the child) persisted less on a subsequent task (Leonard et al., 2021).

In general, prior work has painted a picture of young children as savvy social learners who not only benefit from their own exploration but also from information provided by others (e.g., Gweon, 2021). Importantly, there is rarely a single "best" strategy in learning; both exploration and help-seeking are important and useful, but their effectiveness varies depending on the cost of exploration and the likelihood of receiving help. Some recent work has shown that children preferentially explore aversive stimuli that were learned when their own caregiver was present (Tottenham et al., 2019) and also expect other children to explore more in the presence of their own caregiver (Dahmani et al., 2024), suggesting that children may intuitively understand the benefits of learning in supportive social environments. However, most prior work has presumed a reliable learning environment and thus does not afford the opportunity to show the strengths of children who are adapted to, and assessed in, contexts of unreliability (Ellis et al., 2022). In this work, we demonstrate the double-edged nature of such adaptations while emphasizing the importance of studying learning in contexts that reflect the full range of social environments children may encounter (DeJoseph et al., 2024).

Limitations and guides to interpretation

The current work represents an experimental approach to manipulate the reliability of help in a lab setting. Thus, it is useful to unpack the manipulation in more detail to understand what varied across conditions and its potential limitations. In all of our experiments, the history phase consisted of two consecutive scenarios, both of which involved an experimenter promising help, leaving the participant alone with a stuck object, and then returning after one minute to either offer the needed help (Reliable) or say she was too busy to help and offer a less rewarding alternative (Unreliable). What were the consequences of this manipulation? First, as hypothesized, this manipulation affected children's beliefs about the likelihood of receiving help in the future. Our findings suggest that even though both scenarios involved a relatively narrow task domain—opening a stuck object—children readily generalized from such experiences and inferred the experimenter's likelihood of helping on a different future task, such as a puzzle or a causal toy. Second, this manipulation may also have elicited potential differences in children's affective experiences. For instance, those in the Unreliable condition may have felt disappointment or even frustration, both from being unable to achieve a desired goal and from realizing their bids went unanswered. While this may have affected children's tendency to choose an easier task in Experiment 1, our findings in Experiment 2 suggests that even if children experienced such emotions, they did not simply result in overall disengagement; rather, children who experienced unreliable help persisted in exploration more than those who experienced reliable help. Yet, these emotions are often part of the ecological reality of not having help when needed, which may have important consequences—above and beyond their learning decisions—that we were unable to identify in the current study.

The current work, however, does not directly address the scope of children's generalization: Is the experimenter having a busy day, or is she generally unreliable (situation vs. trait)? Is she unreliable specifically in her availability of help, or is she also more likely to be unreliable in other ways (e.g., ignorant or deceptive)? Did I encounter an unreliable individual, or is this a property of my broader social environment? In this sense, it is important to acknowledge that the current work manipulated just one facet of real-world unreliability—momentary nonresponse despite a prior promise (e.g., an adult who clearly intends to help but is busy)—without being able to measure the exact scope of children's inferences from these experiences. In real-world contexts, children might also experience the unreliability of help in other ways, from momentary unavailability due to physical absence to more chronic forms of unresponsiveness (or even outright neglect), either from a single individual or from multiple adults in their environment. In the current study, we chose a rather transient, light-dose form of unreliability because it is common, ethically implementable, and isolates the causal role of nonresponse to a bid for help. While it is notable that children in our study were sensitive even to this light dose of unreliability, open questions remain regarding how children might respond to other forms of unreliability in the availability of help.

One particularly important consideration for interpreting our results is the relatively high socioeconomic background of our sample: Participants were recruited in a science museum located in an affluent city in the United States. The goal of the current work was not to approximate the complexities of learning in real-world environments, but rather to experimentally manipulate a single factor—the reliability of help—to find empirical evidence of adaptation that aligns with findings from prior work on early learning. In this light, our approach—convenience sampling from the local population—was a reasonable first step for at least two reasons. First, this is consistent with most prior work on children's sensitivity to epistemic reliability (e.g., Harris et al., 2018; Sobel & Kushnir, 2013) as well as other forms of environmental reliability (e.g., Kidd et al., 2013; Michaelson & Munakata, 2016). Second, given that most children may experience variability in the reliability of help regardless of socioeconomic status (SES), there was no a priori scientific reason to constrain our sample based on their demographics. In fact, given that children in our sample presumably expect adults to be generally dependable and trustworthy (Stamos et al., 2019), it is notable that we found evidence of adaptation even to a relatively "light" dose of unreliable help (e.g., two interactions with a generally positive tone). At the same time, it is also possible that the experience of unreliable help was especially salient for these children, contributing to the effectiveness of such a light dose of unreliability. In this sense, the relatively high SES of our sample limits the extent to which we can generalize our findings to populations that might experience more variable, or less supportive, forms of caregiving.

The current findings thus raise important questions about how such adaptation might manifest beyond the laboratory setting. Further work is needed to understand how children calibrate their learning behaviors in response to past social reliability in real-world settings, and how these adaptations may carry forward across new contexts. For example, future research might extend these findings by recruiting participants from more diverse socioeconomic backgrounds, thus likely including greater variation in the extent of social support that children receive, and examining whether the same adaptive mechanisms identified here scale to these contexts (Barocas et al., 1991; Evans, 2004). Of particular importance is studying children who are exposed to more chronic, pervasive forms of unreliability. Children whose caregivers suffer from poverty-related stress are more likely to experience unresponsive or inconsistent caregiving (Evans, 2004; Evans et al., 1999; Pinderhughes et al., 2001; Roeters et al., 2010). The consequences of such unreliability—e.g., low motivation, reduced persistence, or disengagement from learning—are precisely the kinds of behaviors that have traditionally been interpreted through a deficit lens, with the assumption that unresponsive caregiving produces maladaptive behaviors that undermine children's cognitive and academic outcomes (Patterson et al., 1992). Consistent with this idea, studies have shown that children who experience their lives as less predictable tend to explore less for information, because of a preference for familiarity and a tendency to repeat their previous responses (Xu et al., 2023; see also Harms et al., 2018). Thus, future work that specifically targets such a population may offer insights that further challenge and complement prevailing deficit models of development under adversity.

Conclusion

Despite the limitations, the current work offers important insights that can guide parenting and educational decisions in everyday contexts. Our findings suggest that even young children are sensitive to the past reliability of their social partners and strategically guide their own learning. In this light, asking for help on an easy task or insisting on trying an impossible task may not necessarily reflect a failure to learn effectively; instead, these behaviors may reflect an attempt to balance the expected utility of exploration against the likelihood of receiving help. In this sense, supportive caregiving and education are not simply about offering help, but about offering help when needed, and providing informative feedback when children are miscalibrated in terms of task difficulty, the availability of help, or the value of exploration. The current work can also inform efforts to support school readiness. For example, educators and social workers who might initially view a help-avoidant child as unmotivated or antisocial may instead focus on building trust and supporting flexible social learning strategies by helping children learn when to rely more—or less—on others' assistance depending on the context (Landry et al., 2017; Sheridan et al., 2008).

Looking forward, the current work highlights an important direction for future research and practice: reconsidering what it means to provide supportive learning environments in light of the variability in children's past experiences of social reliability. By evaluating children's behavior in light of their prior social reliability and encouraging flexibility in both help-seeking and exploration, educators can help children recalibrate their expectations and make learning decisions that are adaptive within and across different social contexts. In this way, we may better support children to maximize their learning potential, regardless of their prior experiences.

Supplementary material

Supplementary material is available at *Child Development* online.

Data availability

The data and analytic code necessary to reproduce the analyses presented here are publicly accessible on OSF (see main text). The materials necessary to attempt to replicate the findings (e.g., images of stimuli used) are publicly accessible in the main text. The analyses presented here were preregistered; please refer to the links provided in each experiment.

Author contributions

Katherine Shannon (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Validation [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Aneesa Conine-Nakano (Data curation [supporting], Investigation [supporting], Project administration [supporting], Writing—original draft [supporting], Writing—review & editing [supporting]), Willem Frankenhuis (Conceptualization [equal], Funding acquisition [equal], Methodology [equal], Resources [equal], Supervision [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Michael Frank (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Supervision [equal], Validation [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), and Hyowon Gweon (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Supervision [equal], Validation [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal])

Funding

This work was supported by the National Institute of Child Health and Human Development (NICHD; F32HD105705, awarded to KAS); the Dutch Research Council (NWO; Vidi grant V1.Vidi.195.130, awarded to WEF); the James S. McDonnell Foundation (grants awarded to WEF: https://doi.org/10.37717/220020502; and to HG: https://doi.org/10.37717/220020545); the Jacobs Foundation (awarded to MCF); and the National Science Foundation (NSF; BCS-2019567 and CCRI-2120095, awarded to HG).

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgments

We are grateful to the research assistants and participating families who made this work possible.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*, 0064. https://doi.org/10.1038/s41562-017-0064

Barocas, R., Seifer, R., Sameroff, A. J., Andrews, T. A., Croft, R. T., & Ostrow, E. (1991). Social and interpersonal determinants of developmental risk. *Developmental Psychology*, *27*, 479–488. https://doi.org/10.1037/0012-1649.27.3.479

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*, 6–10. https://doi.org/10.1038/s41562-017-0189-z

Birch, S. A.J., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, *107*, 1018–1034. https://doi.org/10.1016/j.cognition.2007.12.008

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *The American Psychologist*, *57*, 111–127. https://doi.org/10.1037/0003-066X.57.2.111

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*, 322–330. https://doi.org/10.1016/j.cognition.2010.10.001

Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2020). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour*, *4*, 144–152. https://doi.org/10.1038/s41562-019-0748-6

Cluver, A., Heyman, G., & Carver, L. J. (2013). Young children selectively seek help when solving problems. *Journal of Experimental Child Psychology*, *115*, 570–578. https://doi.org/10.1016/j.jecp.2012.12.011

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B*, *362*, 933–942. https://doi.org/10.1098/rstb.2007.2098

Cook, K. D., & Coley, R. L. (2017). School transition practices and children's social and academic adjustment in kindergarten. *Journal of Educational Psychology*, *109*, 166–177. https://doi.org/10.1037/edu0000139

Corriveau, K. H., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*, 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x.

Corriveau, K. H., Kinzler, K. D., & Harris, P. L. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology*, *49*, 470–479. https://doi.org/10.1037/a0030604

Dahmani, A., Amir, D., Thomas, A. J., & Gopnik, A. (2024). Caregiver presence promotes judgements of exploration. Proceedings of the Annual Meeting of the Cognitive Science Society, 46. Retrieved from https://escholarship.org/uc/item/6s33x461.

De Castella, K., Byrne, D., & Covington, M. (2013). Unmotivated or motivated to fail? A cross-cultural study of achievement motivation, fear of failure, and student disengagement. *Journal of Educational Psychology*, *105*, 861–880. https://doi.org/10.1037/a0032464

DeJoseph, M. L., Ellwood-Lowe, M. E., Miller-Cotto, D., Silverman, D., Shannon, K. A., Reyes, G., Rakesh, D., & Frankenhuis, W. E. (2024). The promise and pitfalls of a strength-based approach to child poverty and neurocognitive development: Implications for policy. *Developmental Cognitive Neuroscience*, *66*, 101375. https://doi.org/10.1016/j.dcn.2024.101375

Dicintio, M. J., & Gee, S. (1999). Control is the key: Unlocking the motivation of at-risk students. *Psychology in the Schools*, *36*, 231–237.

Ellis, B. J., Abrams, L. S., Masten, A. S., Sternberg, R. J., Tottenham, N., & Frankenhuis, W. E. (2022). Hidden talents in harsh environments. *Development and Psychopathology*, *34*, 95–113. https://doi.org/10.1017/S0954579420000887

Ellis, B. J., Abrams, L. S., Masten, A. S., Sternberg, R. J., Tottenham, N., & Frankenhuis, W. E. (2023). *The hidden talents framework: Implications for science, policy, and practice. Elements in applied evolutionary science*. Cambridge University Press https://doi.org/10.1017/9781009350051

Etikan, I. (2016). Comparison of snowball sampling and sequential sampling technique. *Biometrics & Biostatistics International Journal*, *3*, 55. https://doi.org/10.15406/bbij.2016.03.00055

Evans, G. W. (2004). The environment of childhood poverty. *The American Psychologist*, *59*, 77–92. https://doi.org/10.1037/0003-066X.59.2.77

Evans, G. W., Maxwell, L. E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology*, *35*, 1020–1023. https://doi.org/10.1037/0012-1649.35.4.1020

Frankenhuis, W. E., & Del Giudice, M. (2012). When do adaptive developmental mechanisms yield maladaptive outcomes? *Developmental Psychology*, *48*, 628–642. https://doi.org/10.1037/a0025629

Frankenhuis, W. E., & de Weerth, C. (2013). Does early-life exposure to stress shape or impair cognition? *Current Directions in Psychological Science*, *22*, 407–412. https://doi.org/10.1177/0963721413484324

Frankenhuis, W. E., & Gopnik, A. (2023). Early adversity and the development of explore–exploit tradeoffs. *Trends in Cognitive Sciences*, *27*, 616–630. https://doi.org/10.1016/j.tics.2023.04.001

Frankenhuis, W. E., Nettle, D., & McNamara, J. M. (2018). Echoes of early life: Recent insights from mathematical modeling. *Child Development*, *89*, 1504–1518. https://doi.org/10.1111/cdev.13108

Frankenhuis, W. E., Young, E. S., & Ellis, B. J. (2020). The hidden talents approach: Theoretical and methodological challenges. *Trends in Cognitive Sciences*, *24*, 569–581. https://doi.org/10.1016/j.tics.2020.03.007

Gweon, H. (2021). Inferential social learning: Cognitive foundations of human social learning and teaching. *Trends in Cognitive Sciences*, *25*, 896910. https://doi.org/10.1016/j.tics.2021.07.008

Gweon, H., & Asaba, M. (2018). Order matters: Children's evaluation of underinformative teachers depends on context. *Child development*, *89*, e278–e292. https://doi.org/10.1111/cdev.12825

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

Gweon, H., & Schulz, L. (2011). 16-month-olds rationally infer causes of failed actions. *Science*, *332*, 1524–1524. https://doi.org/10.1126/science.1204493

Harms, M. B., Shannon Bowen, K. E., Hanson, J. L., & Pollak, S. D. (2018). Instrumental learning and cognitive flexibility processes are impaired in children exposed to early life stress. *Developmental Science*, *21*, e12596. https://doi.org/10.1111/desc.12596

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

Hembacher, E., deMayo, B., & Frank, M. C. (2020). Children's social information seeking is sensitive to referential ambiguity. *Child Development*, *91*, e1178–e1193. https://doi.org/10.1111/cdev.13427

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*, 589–604. https://doi.org/10.1016/j.tics.2016.05.011

Jaswal, V. K., & Neely, L. A. (2006). Adults Don't Always Know Best. *Psychological Science*, *17*, 757–758. https://doi.org/10.1111/j.1467-9280.2006.01778.x

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*, 109–114. https://doi.org/10.1016/j.cognition.2012.08.004

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

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

Kuchirko, Y. A., & Tamis-LeMonda, C. S. (2019). The cultural context of infant development: Variability, specificity, and universality. In D. A. Henry, E. Votruba-Drzal, & P. Miller (Eds.), *Advances in child development and behavior* (Vol. *57*, pp. 27–63). JAI. https://doi.org/10.1016/bs.acdb.2019.04.004

Landry, S. H., Zucker, T. A., Williams, J. M., Merz, E. C., Guttentag, C. L., & Taylor, H. B. (2017). Improving school readiness of high-risk preschoolers: Combining high quality instructional strategies with responsive training for teachers and parents. *Early Childhood Research Quarterly*, *40*, 38–51. https://doi.org/10.1016/j.ecresq.2016.12.001

Lee, W. S. C., & Carlson, S. M. (2015). Knowing when to be "rational": Flexible economic decision making and executive function in preschool children. *Child Development*, *86*, 1434–1448. https://doi.org/10.1111/cdev.12401

Leonard, J. A., Martinez, D. N., Dashineau, S. C., Park, A. T., & Mackey, A. P. (2021). Children persist less when adults take over. *Child Development*, *92*, 1325–1336. https://doi.org/10.1111/cdev.13492

Linde, M., & van Ravenzwaaij, D. (2023). baymedr: an R package and web application for the calculation of Bayes factors for superiority, equivalence, and non-inferiority designs. *BMC Medical Research Methodology*, *23*, https://doi.org/10.1186/s12874-023-02097-y

Liquin, E. G., & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, *218*, 104940. https://doi.org/10.1016/j.cognition.2021.104940

Michaelson, L. E., & Munakata, Y. (2016). Trust matters: Seeing how an adult treats another person influences preschoolers' willingness to delay gratification. *Developmental Science*, *19*, 1011–1019. https://doi.org/10.1111/desc.12388

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

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, *16*, 329–337. https://doi.org/10.1037/h0029815

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*, *54*, 687–696. https://doi.org/10.1037/0022-3514.54.4.687

Moffett, L., Flannagan, C., & Shah, P. (2020). The influence of environmental reliability in the marshmallow task: An extension study. *Journal of Experimental Child Psychology*, *194*, 104821. https://doi.org/10.1016/j.jecp.2020.104821

Mulder, J., Williams, D. R., Gu, X., Tomarken, A., Böing-Messing, F., Olsson-Collentine, A., Meijerink-Bosman, M., Menke, J., van Aert, R., Fox, J. P., Hoijtink, H., Rosseel, Y., Wagenmakers, E. J., & van Lissa, C. (2021). BFpack: Flexible Bayes factor testing of scientific theories in R. *Journal of Statistical Software*, *100*, 1–63. https://doi.org/10.18637/jss.v100.i18

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants.

Developmental Psychology, *43*, 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *A social learning approach to family intervention: Antisocial boys* (vol. *4*). Castalia Publishing Company.

Pesch, A., & Koenig, M. A. (2023). Trust Matters: Measuring and Identifying a Role for Epistemic and Interpersonal Trust in Preschoolers' Learning from Teachers. *Early Education and Development*, *34*, 27–52. https://doi.org/10.1080/10409289.2021.1991729

Pinderhughes, E. E., Nix, R., Foster, E. M., & Jones, D. (2001). Parenting in context: Impact of neighborhood poverty, residential stability, public services, social networks, and danger on parental behaviors. *Journal of Marriage and the Family*, *63*, 941–953. https://doi.org/10.1111/j.1741-3737.2001.00941.x

Roeters, A., van der Lippe, T., & Kluwer, E. S. (2010). Work characteristics and parent-child relationship quality: The mediating role of temporal involvement. *Journal of Marriage and the Family*, *72*, 1317–1328. https://doi.org/10.1111/j.1741-3737.2010.00767.x

Ronfard, S., Lane, J. D., Wang, M., & Harris, P. L. (2017). The impact of counter-perceptual testimony on children's categorization after a delay. *Journal of Experimental Child Psychology*, *163*, 151–158. https://doi.org/10.1016/j.jecp.2017.06.006

Scofield, J., & Behrend, D. A. (2008). Learning words from reliable and unreliable speakers. *Cognitive Development*, *23*, 278–290. https://doi.org/10.1016/j.cogdev.2008.01.003

Sheridan, S., Marvin, C., Knoche, L., & Edwards, C. (2008). Getting ready: Promoting school readiness through a relationship-based partnership model. *Early Childhood Services*, *2(3)*, 149–172.

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

Stamos, A., Altsitsiadis, E., & Dewitte, S. (2019). Investigating the effect of childhood socioeconomic background on interpersonal trust: Lower childhood socioeconomic status predicts lower levels of trust. *Personality and Individual Differences*, *145*, 19–25. https://doi.org/10.1016/j.paid.2019.03.011

Stipek, D. (1996). Motivation and instruction. In D. C. Berliner, & R. C. Calfee (Eds.), *Handbook of educational psychology* (Vol. *1*, pp. 85–113). Macmillan.

Suor, J. H., Sturge-Apple, M. L., Davies, P. T., & Cicchetti, D. (2017). A life history approach to delineating how harsh environments and hawk temperament traits differentially shape children's problem-solving skills. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *58*, 902–909. https://doi.org/10.1111/jcpp.12718

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

Therneau, T. M. (2023). A package for survival analysis in R (Version 3.5-7) [R package]. Comprehensive R Archive Network (CRAN). Retrieved March 1, 2024 from https://CRAN.R-project.org/package=survival.

Tottenham, N., Shapiro, M., Flannery, J., Caldera, C., & Sullivan, R. M. (2019). Parental presence switches avoidance to attraction learning in children. *Nature Human Behaviour*, *3*, 1070–1077. https://doi.org/10.1038/s41562-019-0656-9

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š, Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*, 813–826. https://doi.org/10.3758/s13423-020-01798-5

von Stumm, S., & Latham, R. M. (2018). Early life experiences: Meaningful differences within and between families. *Infant Behavior & Development*, *53*, 56–63. https://doi.org/10.1016/j.infbeh.2018.09.001

Xu, F. (2019). Towards a rational constructivist theory of cognitive development. *Psychological Review*, *126*, 841–864. https://doi.org/10.1037/rev0000153

Xu, Y., Harms, M. B., Green, C. S., Wilson, R. C., & Pollak, S. D. (2023). Childhood unpredictability and the development of exploration. *Proceedings of the National Academy of Sciences*, *120*, e2303869120. https://doi.org/10.1073/pnas.2303869120