CHILD DEVELOPMENT

EMPIRICAL ARTICLE

Children Predict Improvement on Novel Skill Learning Tasks

Xiuyuan Zhang | Brandon A. Carrillo | Ariana Christakis | Julia A. Leonard

Department of Psychology, Yale University, New Haven, Connecticut, USA

Correspondence: Xiuyuan Zhang (flora.zhang@yale.edu)

Received: 31 July 2024 | **Revised:** 19 December 2024 | **Accepted:** 19 January 2025

Funding: This study was supported by a Jacobs Foundation Research Fellowship awarded to J.A.L.

Keywords: development | learning | learning curves | predictions of learning

ABSTRACT

Learning takes time: Performance usually starts poorly and improves with practice. Do children intuit this basic phenomenon of skill learning? In preregistered Experiment 1 ($n = 125$; 54% female; 48% White; collected 2022–2023), US 7- to 8-year-old children predicted improved performance, 5- to 6-year-old children predicted flat performance, and 4-year-old children predicted near-instant success followed by worse performance on a novel skill learning task. In preregistered Experiment 2 ($n = 75$; 47% female; 69% White; collected 2023), on a task with lowered cognitive demands, US 4- to 6-year-old children predicted improved performance. Thus, although children expect to improve on novel tasks, younger children need scaffolding to form these predictions and grasp this fundamental aspect of learning.

Picture a child trying to ride a bike for the first time. Do they expect to nail it on the first try? Or instead, do they anticipate that learning to ride a bike will take many attempts and be a process of gradual improvement? Critically, how children think about their future learning could have downstream consequences for their actual learning. For instance, if a child incorrectly assumes that success will come easily, they may pick tasks that are too challenging for their ability, get frustrated, and ultimately give up trying when they don't progress as quickly as expected (Finn and Metcalfe 2008; Dai et al. 2018; Muenks et al. 2018; Leonard et al. 2023). Thus, understanding how children predict their future learning may not only provide insights into children's beliefs about skill acquisition but also inform caregivers on how best to guide children's effort allocation. Here, we explore how young children think their future learning will unfold.

A common way to represent and track learning progress is with a "learning curve"—a measure of performance over time. Learning curves have been well studied and characterized in the domain of motor learning, with decades of research showing that motor skill learning curves follow an exponential decay function (Luft and Buitrago 2005; Heathcote et al. 2000; Krakauer et al. 2019). That is, when learning a new skill, people usually make rapid progress early on, followed by a performance plateau (Luft and Buitrago 2005; Mazur and Hastie 1978). Adults correctly intuit that their future skill learning will have this shape even before engaging in a task (Zhang et al. 2025). Children's performance similarly improves over time on motor tasks (Rossi et al. 2019; Solum et al. 2020). However, less is known about children's understanding of how their future learning unfolds over time. Specifically, it is unclear whether children, like adults, expect that their performance will increase during the beginning of skill acquisition.

Prior work suggests that, with experience, preschool-age children can accurately represent features of their future performance (e.g., expectancy of success, or uncertainty) at discrete time points and use this information to guide their actions. For

Xiuyuan Zhang and Brandon A. Carrillo contributed equally to this study.

example, 20-month-old toddlers selectively ask their caregiver for help when encountering a challenge (Goupil et al. 2016) and preschoolers explore more when they are uncertain (Lapidow et al. 2022; Baer and Kidd 2022; Ghetti et al. 2013; Schulz and Bonawitz 2007). Furthermore, 4- to 6-year-old children are more likely to stick with a challenge when their performance has improved compared to when it has stayed the same, suggesting that children are sensitive to the rate of their past performance (Leonard et al. 2023). However, most of these studies looked at children's behavior after some experience with a task. It remains less clear whether children can predict features of their future learning *before* embarking on a novel task. Furthermore, it is unknown whether children only represent their future learning at one time point or over a sequence of points in time (i.e., a learning curve).

Children's decisions about what to practice provide another clue for discovering how they think about their future learning. Specifically, studies on deliberate practice reveal a developmental change in practice strategies between the ages of 4 and 8. By age 6, children selectively practice the game (amongst other games) that they will soon be tested on (Brinums et al. 2018; Casey and Redshaw 2022). When asked why they decided to play the soon-to-be-tested game, most 6- to 7-year-old children referenced ‘practice’, whereas 4- to 5-year-old children did not (Brinums et al. 2018; see also Davis et al. 2015). Similarly, when asked to describe “what does learning mean?”, 6- to 8-year-old children used more process-based language (e.g., references to practice) than 4- to 5-year-old children (Sobel and Letourneau 2015). Taken together, this body of work shows that by age 6, children seem to understand that practice is necessary for mastery and verbally describe learning as improvement (Brinums et al. 2018; Casey and Redshaw 2022; Davis et al. 2015; Sobel and Letourneau 2015).

This developmental change in children's understanding of practice may be due to an age-related shift in how children represent their future learning. Decades of research have shown that young children are unduly optimistic when making predictions about their future performance (Flavell et al. 1970; Yussen and Levy 1975; Schneider 1998; Coote and Livesey 1999; Hennefield and Markson 2022; Xia et al. 2024; Leonard and Sommerville 2024). For example, even after receiving feedback that their predictions are too lofty, 4-year-old children continue to over-predict their performance on a motor task across trials (Xia et al. 2023; Schneider 1998). This optimism may cause preschool-age children to expect quick success on novel tasks and thus not appreciate the necessity of practice. Although optimism declines between the ages of 3 and 11, 7- to 11-year-old children are still optimistic about their own learning when compared to adults (Lockhart et al. 2021; Xia et al. 2024; Leonard and Sommerville 2024). Thus, even though elementary school-age children appreciate the need for practice (e.g., Brinums et al. 2018; Casey and Redshaw 2022), they may still over-predict parts of their future learning curve.

It is also possible that prior work underestimated young children's abilities to predict their future learning. Many of these studies relied on cognitively demanding paradigms that tax children's verbal abilities, memory, and executive function—all skills that are known to develop with age (Best and Miller 2010; Gathercole 1998; Hunt 1978). Indeed, recent work by Serko and colleagues (Serko et al. 2024) found that, on a simplified deliberate practice paradigm where children were unsure which one of two tasks would be tested, even 4- to 5-year-old children chose to practice the harder of the two tasks after experience with both. Thus, on simplified paradigms, even preschool-age children may understand that mastering difficult tasks takes practice. Yet, prior work does not reveal whether children understand the consequences of the decision to practice (e.g., that practice leads to improved performance over time), nor whether they have a graded understanding of how learning will progress. Moreover, it is unclear what children think about their future learning on novel tasks, where they have little experience or feedback on their ability.

1 | The Current Experiment

Here, we set out to understand the nature and developmental timing of children's beliefs about their future learning curves. In particular, we explore whether children, like adults (Zhang et al. 2025), think that they will improve at a novel task over time. To capture even young children's intuitions, we created a novel bean bag toss paradigm with minimal verbal cues and low memory demands. The goal of the game was to toss bean bags onto a target in the middle of a gridded mat on the floor. To make the game novel and challenging, we had children toss the bean bag with their feet instead of their hands. To index children's anticipated learning curve, we had them predict where their first few tosses would land by physically placing bean bags on the mat. A benefit of capturing predictions of learning in motor skill task, as opposed to a more opaque cognitive task, is that the motor task does not impose additional attentional and processing demands on children (Hiscock et al. 1985; Miller et al. 1991; Schäfer 2005). Instead, capitalizing on children's intuitive understanding of relative distance (Boyer and Levine 2015; de Hevia and Spelke 2010; Huttenlocher et al. 1994; Newcombe 2014), children made non-verbal predictions of future performance using the tangible properties of the task. Additionally, each child's prior predictions were left in place on the mat to reduce any memory demands of past predictions.

Using the grid on the mat, we recorded the coordinates of each bean bag and calculated the Euclidean distance (from the bean bag to the center of the target), which served as our primary measure of performance. With this information, we were able to reconstruct each child's precise predicted learning curve. Distinct from prior work (see Xia et al. 2023; Schneider 1998), children made each prediction without performance feedback in between, allowing us to get an estimate of children's predicted learning curves unbiased by actual performance. After children made predictions, we recorded their actual performance while playing the game as a reference point for examining their optimism.

In Experiment 1, we examined 4- to 8-year-old children's performance predictions, and in Experiment 2, we honed in on 4- to 6-year-old children on an adapted task (Experiment 1 preregistration: https://osf.io/mw5tq; Experiment 2 preregistration: https://osf.io/ux79b; all analyses reported in the Result sections are preregistered unless otherwise noted, see Data S1 for minor

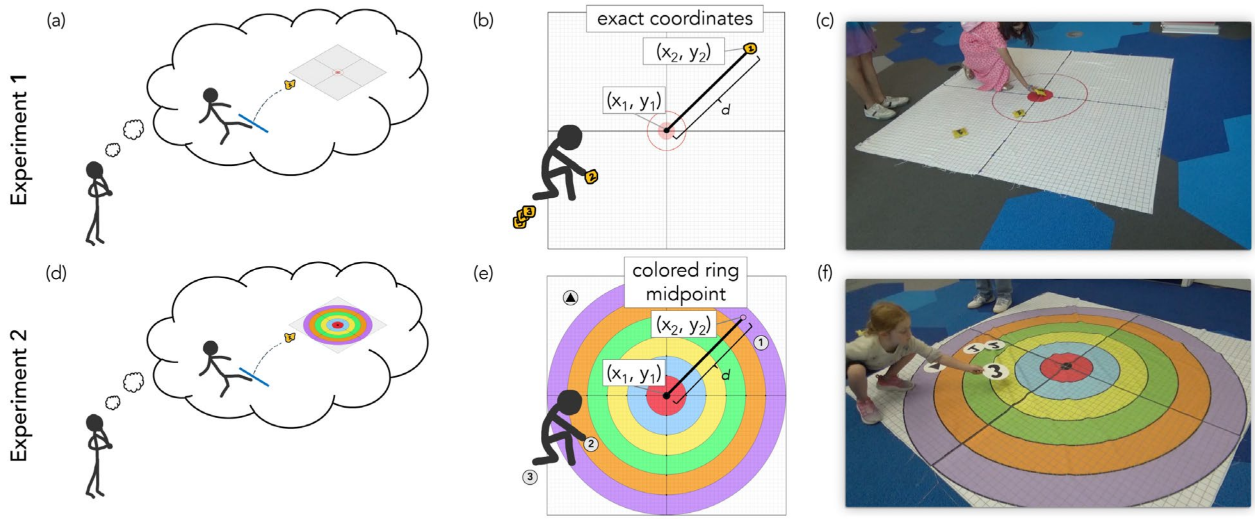

FIGURE 1 | Experiment schematics. (a) In Experiment 1, 4- to 8-year-old children were introduced to a game where they had to toss bean bags (5 total tosses) with their foot to try to land them in a red circle on a gridded mat 8 ft. away. (b) In Experiment 1, before playing, children were asked to make predictions about where they thought each of their tosses would land. They did so by placing the bean bags (with numbers for each toss) directly on the mat. The Euclidean distance (d) between each bean bag's location and the center of the grid was calculated using $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$, in order to reconstruct a child's predicted and actual performance. (c) Photo of participant placing predictions on the mat in Experiment 1. (d) In Experiment 2, 4- to 6-year-old children played a similar game as in Experiment 1 with lower cognitive demands and features to reduce optimism. Specifically the gridded mat now included concentric, colored rings as the prediction space and children only made 3 total bean bag tosses. Children were also given a chance to practice making a prediction before making a practice toss towards the mat, and they were told that they could receive rewards for making correct predictions (bean bag landing in the predicted region). (e) Children made predictions about where they thought each of their tosses would land, this time using numbered paper plates corresponding with the numbered bean bags and placing predictions inside rings. Children were given the option to either place the plates on the mat or say the color of the ring they predicted in Experiment 2. (f) Photo of participant placing predictions on the modified mat in Experiment 2.

deviations from preregistration; OSF repository with script, data, and code: https://osf.io/wtyxn). We focused on this broad age range to capture potential developmental changes given prior work showing maturation in optimism and understanding of practice between the ages of 4 and 8 (Brinums et al. 2018; Casey and Redshaw 2022; Leonard and Sommerville 2024; Xia et al. 2024). We hypothesized that children would predict that their performance would improve across trials (i.e., the bean bags would get closer to the center; note that we did not have predictions about *where* in particular the bean bags would land). However, given young children's documented optimism (Xia et al. 2024; Leonard and Sommerville 2024), we also hypothesized that younger children would predict better performance than older children, and in turn, potentially flatter learning curves. Taken together, this work aims to uncover whether children approach novel tasks thinking that they will get better.

2 | Experiment 1

The goal of Experiment 1 was to explore whether 4- to 8-year-old children predict that their performance will improve on a novel motor task. Children predicted where their first five tosses of a bean bag toss game would land on a large mat with a coordinate grid and a red circle in the center (see Figure 1a–c; see Xia et al. 2023 and Schneider 1998 for work using a similar amount of trials for children's performance predictions). After children made predictions, they played the game for five trials. Children's actual tosses served as a comparison point for evaluating the accuracy of their predictions of performance and slope of improvement, and in turn, their optimism. Additionally, we asked children to report the perceived task difficulty after making predictions to compare verbal to non-verbal (average performance and slope) task assessments.

2.1 | Method

2.1.1 | Participants

We collected a sample of 125 4- to 8-year-old participants (25 children per age group, binned by age in year; 54% female, 43% male, and 3% preferred not to answer) at two local museums in Hartford, Connecticut and Philadelphia, Pennsylvania from July 2022 to November 2023. All participants were from the United States. We did not perform a power analysis a priori because there was no effect size estimate of predicted learning curves. Instead, we based our sample size on related studies that also have children make inferences about their future performance (e.g., Schneider 1998; Serko et al. 2024). Parental income ranged from less than \$4,999 to more than \$200,000 a year, with a median income of \$125,000 (M = \$122,840, SD = \$60,780; 26% of parents had missing demographic data or preferred not to answer). Parental education also ranged from 12 to 20 years (M = 16.26, SD = 2.36; 6% of parents had missing demographic data or preferred not to answer; see Data S1 exploratory for analyses on household income and parental education). The racial and ethnic makeup of the participants was as follows: 48% White, 19% Hispanic or Latino, 17% Asian, 13% Other, 7% Multiracial, 6% Black, 1% American Indian or Alaskan, 1% Native Hawaiian or Pacific Islander, and 7% preferred not to answer. Additional 20 participants were excluded from further analyses based on preregistered criteria: ASD diagnosis ($n = 6$), children failing to complete the full experiment ($n = 5$), children moving their bean bag tosses before they could be properly recorded ($n = 5$), and experimenter error ($n = 4$). One additional participant ($n = 1$) was excluded due to outside interference (e.g., parent or other child interfering with bean bags; note that this exclusion criteria was not preregistered due to an oversight).

2.1.2 | Stimuli

The setup for the bean bag toss game included a gridded mat and five bean bags. The mat was 84.6 × 84.6 in. (with 60 × 60 1.41 in. squares as the grids). On the mat, we painted a circle in red at the center (4.23 in. in radius) and a concentric outer ring in red (16.92 in. in radius; Figure 1b) to reduce children from overly fixating on the center red circle. The five bean bags were all 4 × 4 in. and each weighed 6 oz. The bean bags were labeled with numbers 1 through 5 on both sides.

2.1.3 | Procedure

This experiment had three phases: a training phase, a prediction phase, and a play phase. In the training phase, participants were told that the goal of the game was to land the bean bags in the red circle at the center of the mat (see Figure 1a). Participants were asked to restate the goal of the game as a comprehension question, were corrected if they answered incorrectly, and were excluded if they answered the question incorrectly on the second try (all participants answered this comprehension check correctly). The experimenter next walked the participant to a line taped 8 ft away from the mat and explained that they needed to toss the bean bag from behind the line. However, instead of tossing with their hand, the experimenter explained that they needed to toss the bean bag with their foot in this game. To gain some understanding of the task, participants were allowed to make one practice toss away from the mat and toward the experimenter standing 2 ft from them. The practice toss was intentionally directed away from the mat to avoid anchoring effects on children's predictions of their performance.

In the prediction phase, participants made five sequential predictions about where their bean bag tosses would land. The experimenter asked, "where do you think your first (Trial 1)/next (Trials 2, 3, and 4)/last (Trial 5) toss will land?" and prompted children to walk on the mat and place a bean bag down to mark their prediction (the number on the bean bag matched the trial number). All five bean bags were left on the mat during the prediction phase. The experimenter took a photo of the bean bags on the mat and recorded the coordinates of off-mat tosses with a measuring tape before collecting all the bean bags for the play phase.

To capture whether participants' predicted performance trajectories related to their explicit perception of how hard the task was going to be (e.g., steeper curves for people who thought the task was easier), we asked participants after their predictions, "do you think landing the bean bag in the center of the mat with your foot is easy or hard?" and a follow-up question, "do you think it is kind of easy (hard) or really easy (hard)?"

In the play phase, participants stood 8 ft away from the mat and made five sequential tosses (the number on the bean bag also matched the trial number). After a participant completed all tosses, the experimenter took a photo of the bean bags on the mat and recorded the coordinates of off-mat tosses with a measuring tape.

2.1.4 | Data Coding

The horizontal and vertical locations for the center of the bean bags were coded based on photos of the mat. If a bean bag landed outside the mat (6% of predicted tosses and 72% of actual tosses), measurements were taken by the experimenter at the time of testing using a measuring tape (from the center of the bean bag to the edge of the grid) and converted to *X*- and *Y*-coordinate values. All *X*- and *Y*-coordinate data were double-scored by a second coder based on photographs and videos (note that bean bags that landed off the mat could not be double-scored; see OSF repository for response coding instructions). A third coder arbitrated discrepancies over 1 in. between the two coders (note that the bean bags were 4 by 4 in., much larger than the grid units of 1.41 by 1.41 in., and thus we allowed discrepancies under 1 in. in measurement). Coder scores were highly correlated ($r = 0.99$, $p < 0.001$) before third-coder arbitration.

We calculated the Euclidean distance based on the *X*- and *Y*-axis locations of the bean bags to the center of the mat. Predicted performance was operationalized as the Euclidean distance from the center of the bean bag to the center of the mat during the prediction phase, and actual performance was operationalized as the Euclidean distance from the center of the bean bag to the center of the mat during the play phase.

2.2 | Results

2.2.1 | Performance Predictions

Contrary to our hypothesis, on average 4- to 8-year-old children did not predict that their performance would improve across trials (Figure 2a). Instead, only children 79 months (6.6 years) and older predicted improved performance across trials. A linear mixed-effects model predicting children's predicted performance (in Euclidean distance) using trial and age (in months) with random intercepts and random slopes of trial by participants revealed a main effect of age ($b = 0.21$, $p < 0.002$) and only a trend of trial ($b = -0.74$, $p = 0.12$; an additional model testing if children's predicted performance followed a quadratic function of trial was not significant, $b = -0.74$, $p = 0.12$, see Data S1 for details). This suggests that, on average, children did not think that their performance would change across trials, but, as hypothesized, younger children predicted better average performance than older children. We also found a significant interaction between trial and age ($b = -0.12$, $p < 0.001$), showing that predictions of future performance change with age. To interrogate age-specific

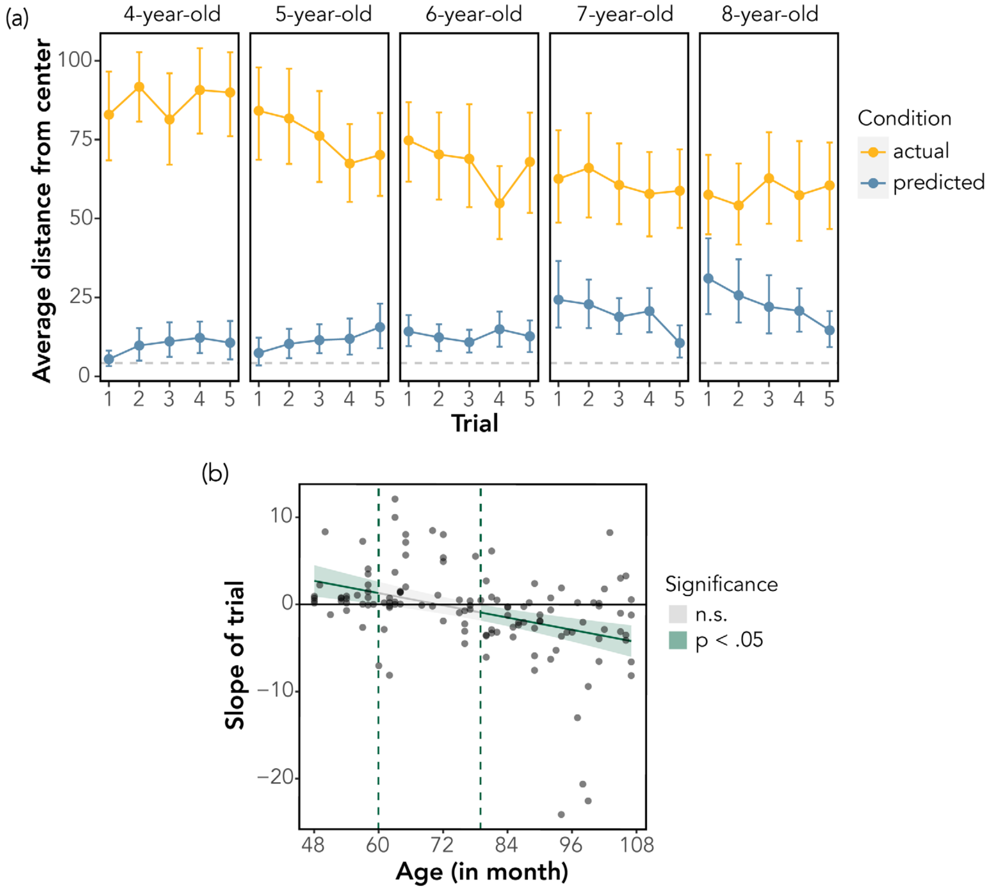

FIGURE 2 | Predicted and actual performance by age in Experiment 1. (a) Average predicted performance (in blue) and actual performance (in yellow) across five trials by age group. Error bars are bootstrapped 95% confidence intervals (computed from simulation with 1000 iterations), and the dashed gray line indicates the radius for the red circle (the center of the mat). A figure with only average predicted performance is available in Data S1. (b) A Johnson-Neyman test showing predicted learning curves over five trials (i.e., slope) by participants' age (in month). Significant slopes are in green and separated by the vertical green dashed lines. The p-value is FDR-corrected. Individual participants' predicted slopes are shown on the plot as scatter points. The shaded green and gray error bars are 95% confidence intervals calculated in the Johnson-Neyman function. Children younger than 60 months predicted significant positive slopes (vertical dashed line on the left), and children older than 79 months predicted significant negative slopes (vertical dashed line on the right).

effects on performance prediction, we conducted a Johnson-Neyman test. As shown in Figure 2b, we found that participants younger than 60 months predicted worse performance over time (i.e., a positive slope of trial; to calculate each participant's slope, we took the beta coefficient, for each participant from the linear regression: predicted distance $= b$ * trial + a), while participants 79 months and older predicted improved performance (i.e., a negative slope, $t = 2.13$, $p_{FDR\text{-}corrected} < 0.05$). Additional exploratory analyses examining the location of children's predictions are included in Data S1.

2.2.2 | Perceived Difficulty (Exploratory)

Younger children thought the task would be easier than older children: A linear regression predicting children's difficulty rating (with a 1-on-1 mapping from "really easy", "kind of easy", "kind of hard", and "really hard" to numeric values 1, 2, 3, and 4) using age revealed a main positive effect of age ($b = 0.02$, $p = 0.004$; $n = 124$ since one participant's difficulty judgment was not available due to experimenter error). Furthermore, controlling for age and trial, children who thought the task would be harder predicted worse performance ($b_{difficulty} = 2.56$, $p = 0.02$; $b_{age} = 0.17$, $p = 0.01$; $b_{trial} = -0.76$, $p = 0.11$ from a linear mixed-effects model predicting children's predicted performance using difficulty rating, age, and trial with random intercepts and slopes of trial by participants). However, children who thought the task would be harder did not predict steeper slopes ($b_{difficulty} = -0.27$, $p = 0.53$; $b_{age} = -0.11$, $p < 0.001$; from a linear model predicting participants' predicted rate of performance using difficulty ratings and age).

2.2.3 | Actual Performance

On average, children did not significantly improve at the game across five trials, and older children tossed the bean bag closer to the target than younger children: A linear mixed-effects model predicting actual performance using trial and age with random intercepts and slopes of trial by participants revealed a main effect of age ($b = -0.57$, $p < 0.001$) but not trial ($b = -1.30$, $p = 0.17$). There was also no significant trial by age interaction on actual performance ($b = 0.01$, $p = 0.80$), showing that children's actual rate of learning did not differ by age.

2.2.4 | Comparison Between Predicted and Actual Performance

Next, we examined whether children were accurate at predicting their future performance. In general, children's average predictions ($M = 15.30$, $SD = 17.49$) were much lower than their average performance ($M = 70.04$, $SD = 36.71$), showing that children overestimated how well they would perform in the task (lower indicates better performance; paired Wilcoxon test $V = 4951$, $r = 0.82$, $p < 0.001$; see Figure 2a). To probe whether the trial or children's age was related to accuracy in children's predictions, we computed a difference score between children's average predicted and actual performance across five trials and ran an exploratory analysis with a linear mixed-effects model predicting the difference scores using age with random intercepts by participants. Model results revealed a main positive effect of age ($b = 0.86$, $p < 0.001$), showing that older children made more accurate predictions about their average performance than younger children (a model with an additional fixed effect and random slopes for trial did not converge, thus as preregistered, we reported the maximally converged model instead; see Data S1 for details and additional analyses). A linear mixed-effects model predicting children's actual performance using predicted performance, age, and trial (with random intercepts and random slopes of trial and predicted performance by participants) revealed that children's predictions of their performance did not scale with their actual performance ($b = 0.09$, $p = 0.37$).

A set of exploratory analyses further examined whether children accurately predict the slope of their performance curves. To calculate the slope of actual performance, we extracted the beta coefficient from the linear regression: actual distance = b * trial + a for each participant. Children's predicted rates of improvement ($M = -0.74$, $SD = 5.23$) were not significantly different from their actual rate ($M = -1.30$, $SD = 10.41$; paired Wilcoxon test $V = 4288$, $r = 0.08$, $p = 0.39$). We next computed a difference score between children's predicted and actual learning rates and explored if children's accuracy in predicting their learning rates changed with age. We ran a linear model predicting the difference score using age and found that older children overestimated their rates of improvement more compared to younger children ($b = -0.13$, $p = 0.02$). Additionally, a linear model predicting children's actual slope using their predicted slope while controlling for age revealed a trend in which children who predicted steeper learning curves improved more across five trials ($b_{trial} = 0.37$, $p = 0.057$; $b_{age} = 0.06$, $p = 0.32$; a structurally-identical model excluding outliers above and below three standard deviations from the mean resulted in similar findings: $b_{trial} = 0.47$, $p = 0.066$; $b_{age} = 0.06$, $p = 0.31$).

2.3 | Interim Discussion

Results from Experiment 1 revealed that children ages 6.6 to 8 years predicted that they would improve at a novel motor task, while 4-year-old children predicted that they would actually get worse at a novel motor task over time, and 5- to 6.6-year-old children predicted flat learning curves. Children of all ages did not significantly improve at the game. Older children perceived the task as more difficult and predicted a more accurate (yet still optimistic) overall performance but steeper slopes of learning than in reality compared to younger children.

It is possible that older children's over-optimism in their predicted performance and rate of improvement may be an artifact of the testing environment: Children may not have significantly improved on the task across 5 trials because they were tested on the museum floor, a distracting environment, and were not incentivized to actually succeed. We tested this hypothesis in Data S1 Experiment 1b and found that 7- to 8-year-old children still do not significantly improve on the first five trials of the bean bag toss game *even* under motivating (children were offered rewards for accuracy) and quieter conditions (children were tested in a private room off the museum floor; see method and results in Data S1). In short, we did not find significant performance improvement on this task across two experiments with different features that often improve performance (such as making performance predictions and offering rewards for performance; Cottini et al. 2018; Xia et al. 2023). This suggests that older children's optimism about their future performance is not an artifact of the testing environment.

Instead, it is likely that this task requires longer periods of practice for significant improvement. Indeed, prior work found similar learning curves for 10-, 18-, and 40-year-old participants on a dart throwing task with their non-dominant hand across 200 trials (Solum et al. 2020; notably, 10-year-old participants had more variable performance compared to 18- and 40-year-old participants). Thus, 6.6- to 8-year-old children in our study were somewhat correct in predicting improvement over time but were overly optimistic about the timescale of this improvement. Five- to 6.6-year-old children, on the other hand, were overly optimistic about their average performance but somewhat accurate in predicting a flat initial learning curve, and 4-year-old children were neither accurate about their actual performance nor rate of improvement, as they predicted systematically getting worse across trials while their actual performance was flat.

However, it is unclear whether 4-year-old children genuinely believe they will get worse on our task and 5- and 6-year-old children think their performance will remain stagnant. Although it could be the case that cognitive changes in development concerning children's metacognition (Gonzales et al. 2022; Schneider 2008; Roebers 2017) or counterfactual reasoning (Rafetseder et al. 2013; Nyhout and Ganea 2019; Kominsky et al. 2021) spurred these results, there are a number of reasons to believe that the developmental changes we are observing are driven instead by task demands. First, many (55% of children in the age range between 4-year-old and 6.6-year-old) put their first bean bag on the target. As this was the only colored portion of the mat, the saliency of this target could have distracted younger children from thinking about the full range of their predictions. For example, some younger children placed all their bean bags in a straight line or in each quadrant of the target. This behavior could reflect a bias towards salient visual features of the mat rather than the goal of the game *or* a lack of understanding of how the placement of these bean bags corresponds to predictions of performance. Second, although we tried to reduce cognitive demands in our paradigm, it is possible that our current experimental design was too conceptually challenging for younger children. Asking children to make precise predictions on a mat with many possibilities (i.e., every point on the 60 × 60 gridded space) might have been quite complex for young children. Third, the fact that many young children predicted immediate success and said that the game would be "easy" suggests that perhaps their overoptimism may have blocked them from considering the necessity of practice. In Experiment 2, we explore whether even 4- to 6-year-old children can predict that their performance will improve on a modified version of this task with lower cognitive demands and features to reduce optimism.

3 | Experiment 2

The goal of Experiment 2 was to see if 4- to 6-year-old children predict that they will get better at a simplified version of the bean bag toss game used in Experiment 1. We made the following changes to the bean bag toss paradigm to lower cognitive demands and temper children's optimism. First, we created a new target that had five colored rings around the main red circle. By adding more color to the mat, we reduced the saliency of the center red circle to encourage children to attend to the full mat space (see Gaspelin et al. 2015; Turoman et al. 2021). The colored rings also constrained the prediction space to accommodate children's developing spatial and counterfactual reasoning (Vasilyeva and Lourenco 2012; Leahy et al. 2022): Instead of having children predict the precise location of their tosses on a large mat as in Experiment 1, we simply had them predict which ring their toss would land in. To further lower cognitive demands, we reduced the number of prediction trials from five to three (as in Schneider 1998). Finally, to minimize children's optimism, we made the following two changes. First, considering prior work showing that practice trials improve children's task performance (Setoh et al. 2016), we allowed children to have one practice prediction and actual toss to give them first-hand experience that the task might be harder than they expected. Second, inspired by work showing that incentivized predictions lower young children's optimism (Xia et al. 2023), we told children that they could get stickers for both accurate predictions and good performance.

Given this simplified paradigm, we hypothesized that all children would predict improved performance across three trials. Additionally, following the results in Experiment 1, we hypothesized that 6-year-old children would predict overall worse performance than 4-year-old children and, in turn, steeper learning curves.

3.1 | Method

3.1.1 | Participants

We collected a sample of 75 4- to 6-year-old participants (25 per age group as in Experiment 1, binned by age in year; 47% female, 53% male) at the same local museum and two local elementary schools in Connecticut from July 2023 to November 2023. A power analysis indicated that 75 participants (25 participants per age group) were required for a power of 0.80 (see Data S1). Parental income ranged from around \$5000 to more than \$200,000 a year, with a median income of \$175,000 (M = \$156,960, SD = \$53,260; 16% of parents had missing demographic data or preferred not to answer). Parental education also ranged from 12 to 20 years (M = 17.38, SD = 2.18; 5% of parents had missing demographic data or preferred not to answer; see Data S1 for exploratory analyses on household income and parental education). The racial and ethnic makeup of the participants was as follows: 69% White, 16% Hispanic or Latino, 9% Asian, 9% Multiracial, 5% Other, 3% Black, 1% Native Hawaiian or Pacific Islander, and 3% preferred not to answer. Additional 15 participants were excluded from further analyses based on preregistered criteria: incorrectly answering comprehension checks (n = 8), ASD diagnoses (n = 4), child failing to complete full experiment (n = 2), or experimenter error (n = 1).

3.1.2 | Stimuli

We created a modified version of the gridded mat in Experiment 1, with concentric colored rings and a red circle at the center of the mat (see Figure 1d–f). All colored rings were designed to be a uniform distance apart (6.7 in.); however, due to fabric printing disfiguration, the six colored rings had about a 0.3 in. range of distance from one another (approximately 6.5–6.8 in. distance between each ring). This experiment included four bean bags (same dimensions and weight as in Experiment 1): one for the practice trial and labeled with a solid triangle shape, and the remaining three bean bags were labeled 1 through 3. Additionally, four paper plates, each with a diameter of 6 in. were similarly labeled: one plate was labeled with a solid black triangle, and the remaining three were labeled 1 through 3. Paper plates were used to make predictions instead of bean bags (as in Experiment 1) so that participants could easily compare their prediction to their actual toss in the practice round.

3.1.3 | Procedure

Experiment 2 contained the same three phases as Experiment 1: training, prediction, and play. In the training phase, participants were told that they could earn one sticker if their bean bag toss landed in the red circle and one sticker if they accurately predicted which color ring their bean bag toss would land in (and thus two stickers if they predicted that their bean bag would land in the red circle and it actually landed there). To make sure participants understood these rules, we provided them with three examples followed by comprehension check questions (see Data S1). As preregistered, participants were excluded if they answered at least one comprehension question incorrectly after the second try (n = 8).

As in Experiment 1, participants were then told how to toss the bean bags using one of their feet. However, different from Experiment 1, children were now given a practice trial to make a prediction on the target and toss the bean bag once using the "triangle" plate and bean bag. The experimenter told participants that they could make their prediction either by placing the plate somewhere on the mat or by telling the experimenter the color of the predicted ring. We included both of these prediction options to encourage thoughtful predictions from children who were shy with their movements (e.g., placing the materials) or their words (e.g., saying the color). Participants then proceeded to make their practice toss and, afterward, were informed of whether or not they earned a sticker and why (i.e., the experimenter pointed out whether the prediction and actual toss were in the same color or in different colors; $n = 4$ of participants received a sticker here). Both the participant's prediction and practice toss were left on the mat for the remainder of the procedure as a reminder of past performance.

In the prediction phase, participants made three sequential predictions about where their actual tosses would land before they played. To encourage children to be thoughtful with their predictions, the experimenter asked participants why they thought their toss would land in the spot they guessed after each prediction. As in Experiment 1, participants were asked if they thought landing a bean bag in the center of the target would be easy or hard, and the degree of the difficulty level (e.g., "really easy", "really hard", "kind of easy", or "kind of hard") after their predictions to look at the connection between perceived task difficulty and predicted learning curves.

Finally, children played the game. However, since children were now incentivized to land their tosses in their predicted rings, rather than just the target, children's actual toss data are less interpretable as a measure of true learning than the comparable data in Experiment 1. For this reason, we only used comparisons between predicted and actual performance as a measure of optimism in performance (average error) rather than the slope of performance. After completing all three tosses, participants collected any stickers they won and were offered a prize for their participation. The experimenter took a photo of the prediction and actual tosses and recorded predictions and bean bag tosses off the mat using a measuring tape.

3.1.4 | Data Coding

All sessions were video and audio recorded with a mic either clipped to the experimenter's shirt or on a table or chair off to the side for ease of coding the child's responses to questions. If children's predictions and actual tosses landed on a colored region, the color responses were recorded and later transformed into a binned distance to the center of the mat. In order to assign a numerical value for the distance between each color ring and the center of the target, the Euclidean distance from the center of the target to the midpoint of each colored ring was calculated and used (i.e., red ring distance = 3.35 in.; blue ring distance = 10.05 in.; yellow ring distance = 16.75 in.; green ring distance = 23.45 in.; orange ring distance = 30.15 in.; purple ring distance = 36.85 in.). In cases where a participant's bean bag toss or prediction plate was not on a colored ring (in the corners or off the mat; 3% of predictions and 78% of the actual performance tosses), the experimenter recorded the *X*- and *Y*-coordinate on the mat or measured the distance in inches from the center of the bean bag (the plate) to the edge of the mat using a tape measure (and converted it to *X*- and *Y*-coordinate values; like in Experiment 1).

The locations for all of the participants' predicted, and actual tosses on the gridded mat were coded by two separate coders, who each recorded either the color of the ring (note that, similar to Experiment 1, bean bags that landed off the mat could not be double-scored). A third coder arbitrated any discrepancies between the two, with the two coders having a high agreement rating before arbitration ($r = 0.99$, $p < 0.001$).

3.2 | Results

3.2.1 | Performance Predictions

As hypothesized, children predicted that their performance would improve across three trials (Figure 3). However, contrary to our hypothesis, we did not find an effect of age on children's predicted average performance or slope. A linear mixed-effects model predicting children's predicted performance using trial and age with random intercepts and random slopes for trial by participants revealed a main effect of trial ($b = -3.81$, $p = 0.002$) but not age ($b = 0.11$, $p = 0.32$). Additionally, a linear mixed-effects model revealed a trend for trial by age interaction ($b = -0.22$, $p = 0.06$).

To examine if children's experience of a practice trial influenced their subsequent predictions, we conducted three structurally similar linear mixed-effects models predicting performance predictions with children's practice prediction, actual practice performance, and the difference between prediction and actual performance, controlling for effects of trial and age, with random intercepts and random slopes by participants. All three models revealed null effects (see Data S1), suggesting that the practice trial did not relate to children's subsequent performance predictions.

As an exploratory analysis, we examined if children strategically made "floor" predictions (predicting their bean bag would land

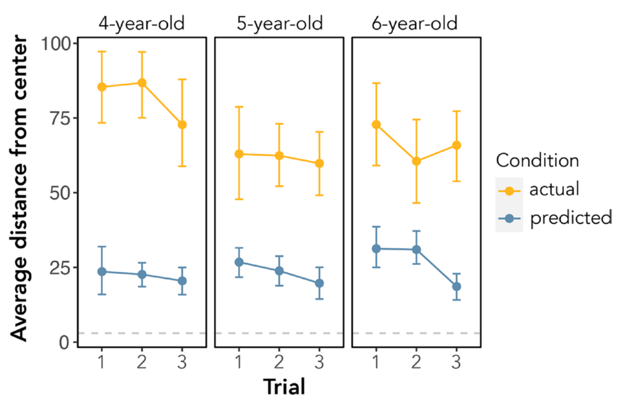

FIGURE 3 | Predicted and actual performance by age in Experiment 2. Average actual (in yellow) and predicted (in blue) performance across three trials by age group. Note that error bars are bootstrapped 95% confidence intervals (computed from simulation with 1000 iterations), and the dashed gray line indicates the radius for the red circle (the center of the mat). A figure with only average predicted performance is available in Data S1.

on the outermost ring or off the mat) to optimize their chances of getting a sticker in the prediction phase. Only two participants placed all their predictions in the outermost ring and one participant placed all their bean bags off the mat. The majority of children placed their predictions on different regions across three trials (9% placed predictions on two regions and 80% placed predictions on three different regions). Thus, children did not try to “hack” the incentive structure of the task to optimize getting stickers and instead presumably tried to accurately predict their future performance.

3.2.2 | Perceived Difficulty (Exploratory)

As in Experiment 1, older children were more likely than younger children to think that the game would be hard (linear regression: $b = 0.03$, $p = 0.01$; $n = 74$ since one participant's difficulty rating was not available due to experimenter error). To interrogate the relation between difficulty judgments and children's performance predictions, we ran a linear mixed-effects model predicting children's predicted performance using their difficulty judgments, controlling for age and trial with random intercepts and random slopes of trial by participants. Children who judged the task as harder also predicted worse performance ($b_{difficulty} = 2.85$, $p = 0.02$; $b_{age} = 0.04$, $p = 0.72$; $b_{trial} = -3.78$, $p = 0.003$). Additionally, consistent with results in Experiment 1, a linear regression predicting children's predicted rate of improvement using difficulty ratings and controlling for age showed no significant relation between difficulty ratings and predicted rate of improvement ($b_{difficulty} = -1.50$, $p = 0.27$; $b_{age} = -0.19$, $p = 0.13$).

3.2.3 | Comparison Between Predicted and Actual Performance (Exploratory)

Since children were incentivized to land their tosses in their predicted rings, their actual performance was bounded by their predicted performance and thus not a good measure of learning. However, comparing average predicted performance to average actual performance can give us a measure of children's optimism. We found that children overestimated their overall performance ($M = 24.21$, $\mathrm{SD} = 14.55$) compared to their actual performance ($M = 69.91$, $\mathrm{SD} = 33.64$; paired Wilcoxon test $V = 696$, $r = 0.80$, $p < 0.001$; see Figure 3). To see whether children's age related to their over-predictions, we conducted a linear mixed-effects model predicting the difference score using age with random intercepts by participants and by trial. Model results revealed a significant effect of age ($b = 0.86$, $p = 0.002$), replicating results from Experiment 1 showing that younger children are more optimistic about their average future performance than older children.

3.2.4 | Comparison Between Actual Performance in Experiments 1 and 2 (Exploratory)

To test whether incentivized performance predictions influenced actual performance, we compared 4- to 6-year-old children's performance on trials 1–3 in Experiment 1, where there was no incentive tied to accurate predictions, to performance in Experiment 2. We found no significant difference between 4- to 6-year-old children's performance in either experiment (Exp. 1: $M = 75.71$, $\mathrm{SD} = 36.10$; Exp. 2: $M = 69.91$, $\mathrm{SD} = 33.64$; two-samples Wilcox: $W = 3104$, $p = 0.27$). Thus, incentivizing accurate performance predictions may not have had a strong influence on children's actual performance in this task.

3.3 | Interim Discussion

Results from Experiment 2 revealed that 4- to 6-year-old children predict that they will improve at a novel motor task on a paradigm with minimal cognitive demands and task features to rein in optimism. Contrary to our hypothesis, we did not find age differences in children's predicted learning curves. However, consistent with results from Experiment 1, we found that younger children over-predicted their actual performance more than older children. Importantly, children did not try to hack the reward system by predicting floor performance and children's overall actual performance did not appear to be biased by the addition of incentives. Taken together, results from Experiment 2 suggest that when tasks are set up to be minimally demanding and explicitly designed to reduce optimism, 4- to 6-year-old children predict that they will improve.

4 | General Discussion

We asked whether 4- to 8-year-old children understand a basic phenomenon of learning: Performance usually improves at the beginning of skill acquisition (Evans et al. 2018; Heathcote et al. 2000). Across two preregistered experiments, we found evidence that children do indeed predict that their performance will get better on the first few trials of a novel skill learning motor task.

Our results reveal that children understand that learning takes time at a younger age than previously realized. In Experiment 1, we found that 6.6- to 8-year-old children, but not younger children, predicted that their performance would get better on a novel task. This finding is in line with prior work showing that by around age 6, children both implicitly and explicitly associate practice with improvement (Brinums et al. 2018; Casey and Redshaw 2022; Davis et al. 2015; Sobel and Letourneau 2015). Going beyond prior work, we found that when we lowered task demands by constraining and familiarizing children with the prediction space and keeping children's optimism at bay, even 4- and 6-year-old children predicted that their performance would improve. However, as we provided ample scaffolding for young children to express their cognitive capacities in Experiment 2 (e.g., changing the visual cues, constraining the prediction space, adding a practice trial, and incentivizing correct guesses), it is unclear which modifications were most impactful. Thus, future work is necessary to tease apart which specific features promote young children's critical reasoning about their future performance.

Our findings also reveal precisely how children think practice will lead to mastery: through gradual improvement in performance. Prior work on deliberate practice shows that 4- to 8-year-old children focus their practice on the soon-to-be-tested game (Brinums et al. 2018), and, when they don't know which game they will be tested on, they opt to practice the harder of two games (Serko et al. 2024). These choices show that children understand that practice is necessary for strong performance but do not reveal exactly *how* children expect to improve with practice. By capturing children's trial-by-trial predictions of their future performance on a task with minimal experience, we show that children do have beliefs about the shape of their future learning curve. In Experiment 1, we find that 6.6- to 8-year-old children expect to gradually, linearly improve at a novel task, and, in Experiment 2, even 4- to 6-year-old children similarly expect incremental improvement. Due to children's limited attention span and cognitive capacities, we only asked children to predict a few trials of learning. In future work, it would be interesting to explore if children, like adults (Zhang et al. 2025), predict the exponential decay shape of motor learning on longer time scales.

Although children's predictions of getting better across trials reflect the improvement seen in average skill acquisition (Evans et al. 2018; Heathcote et al. 2000; Solum et al. 2020), children were not accurate in predicting the specific parameters of their learning curves. First, children of all ages over-predicted their average performance. Across both experiments, we found that younger children were more optimistic about their average performance than older children (in line with work showing that optimism declines with age; see Bamford and Lagattuta 2020; Lockhart et al. 2002; Schneider 1998; Xia et al. 2024; Leonard and Sommerville 2024). Second, 7- to 8-year-old children over-predicted their rate of improvement. It is possible that the size of the mat artificially constrained the prediction space, inflating children's optimism as most of their tosses landed outside the mat. Future research should test whether larger mats that better capture children's true performance would lead to less optimistic predictions. Nonetheless, even on the same prediction mat, younger children still made more optimistic predictions about their future performance than older children, suggesting that age-related changes in over-predictions of performance are not solely driven by the mat size.

Critically, children's optimistic beliefs about their future performance might impact which tasks they choose to pursue and whether they decide to persist or quit them. Prior work shows that children and adults prefer to work on tasks in which their performance has improved (Leonard et al. 2023; Ten et al. 2021). Thus, children may elect to take on tasks that they think they can swiftly master. Our study reveals that young children may need extra scaffolding to predict that learning takes time. Without this scaffolding, young children's optimism about their future learning curve may lead them to pick out tasks that are too hard for them and, in turn, get frustrated and give up when they don't learn as quickly as they thought. Similarly, 7- to 8-year-olds' over-optimism about their rate of improvement may lead to disappointment and potentially quitting when they do not improve as quickly as expected (Zhang et al. 2025; Dai et al. 2018). Thus, measuring and understanding children's predicted learning curves may be useful in educational interventions, allowing educators to catch students' miscalibrations before those errors have serious consequences. A fruitful avenue of future research involves exploring whether and how predictions of learning curves impact children's task choice and persistence.

Our work also adds to broader motivational theories. Specifically, classic work by Carol Dweck and colleagues shows that children with a growth mindset, who think that abilities can grow with practice, are more likely to embrace challenges, persist through them, and have stronger academic outcomes than children with fixed mindsets, who believe talent is unchangeable (Dweck 2006; Haimovitz and Dweck 2017; Dweck and Yeager 2019; Yeager et al. 2019). By probing children's trial-by-trial predictions of their future performance over time, our research reveals exactly *how* children think practice leads to skill improvement. This approach also offers a new, more granular method for measuring children's mindsets (see also Muradoglu et al. 2024). For example, by measuring children's trial-by-trial predictions of future performance on various tasks, we can see whether children's growth mindset on a task correlates with their predicted rate of learning, which opens new doors for research and theory on growth mindset.

Our work has a number of limitations. First, our sample included only participants from a Western, educated, industrialized, rich, and democratic (WEIRD; Henrich et al. 2010) country; thus, we cannot address how our findings might generalize to other populations. Second, we only tested children's predicted learning curves on a motor skill learning task. As such, it is unknown whether children similarly predict that they will improve across all tasks as well as whether they are sensitive to task-specific features that merit this inference. For example, activities that are harder to master generally have flatter learning curves than tasks that are easier to learn (Gottlieb and Oudeyer 2018; Son and Sethi 2006). Although young children are able to detect task difficulty from surface features (Gweon et al. 2017), it is unknown how children's difficulty judgments are related to their predicted learning curves. Our work found that children only link explicit ratings of task difficulty to worse average performance, not flatter slopes. However, it is possible that a more coherent relation between task difficulty and slope of progress would be found if we queried children's prediction of difficulty and task performance on more trials or on a different type of task. Third, it is possible that children's predictions in Experiment 2 reflect some sort of goal setting (e.g., to make the game more enjoyable by setting harder challenges) rather than or, in addition to, their beliefs about their future performance. Future work should see how different explicit goals impact children's predictions of future performance over trials (see Rule et al. 2023).

Children's daily life is marked by learning. Here, we show that young children expect learning to take time: 4- to 8-year-old children predict that their performance will improve with practice on a novel skill learning task. Importantly, 4- to 6-year-old children only make this prediction with added scaffolding. Thus, caregivers and teachers may need to help young children understand this basic phenomenon of learning in order to guide them to take on and persist with optimal challenges.

Acknowledgments

We thank Mika Asaba, Lisa Feigenson, Frank Keil, Melissa Santos, and Reut Shachnai for their feedback and discussions, and research assistance from Saif Behairy, Paloma Casanova, Lizbeth Lozano, Lauren Okine, Yagmur Ozturkoglu, and Natalie Masetti. We also thank members of the Yale Leonard Learning Lab and Yale Cognition and Development Lab for helpful discussions and feedback. Special thanks go to the museums (Connecticut Science Center and Please Touch Museum) and preschools for their partnerships and to the families who participated in this research.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data, analysis code, materials, and preregistrations are publicly accessible on Open Science Framework (OSF): https://osf.io/wtyxn.

References

Baer, C., and C. Kidd. 2022. "Learning With Certainty in Childhood." *Trends in Cognitive Sciences* 26: 887–896. https://doi.org/10.1016/j.tics.2022.07.010.

Bamford, C., and K. H. Lagattuta. 2020. "Optimism and Wishful Thinking: Consistency Across Populations in Children's Expectations for the Future." *Child Development* 91: 1116–1134. https://doi.org/10.1111/cdev.13293.

Best, J. R., and P. H. Miller. 2010. "A Developmental Perspective on Executive Function." *Child Development* 81: 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x.

Boyer, T. W., and S. C. Levine. 2015. "Prompting Children to Reason Proportionally: Processing Discrete Units as Continuous Amounts." *Developmental Psychology* 51: 615–620. https://doi.org/10.1037/a0039010.

Brinums, M., K. Imuta, and T. Suddendorf. 2018. "Practicing for the Future: Deliberate Practice in Early Childhood." *Child Development* 89: 2051–2058. https://doi.org/10.1111/cdev.12938.

Casey, C. G., and J. Redshaw. 2022. "Developmental Changes in Children's Choices of Practice Strategies." *Learning and Instruction* 80: 101625. https://doi.org/10.1016/j.learninstruc.2022.101625.

Coote, K., and D. Livesey. 1999. "Optimism Bias in Children's Motor Performance Expectations." *Australian Educational and Developmental Psychologist* 16: 52–61. https://doi.org/10.1017/S0816512200027267.

Cottini, M., D. Basso, and P. Palladino. 2018. "The Role of Declarative and Procedural Metamemory in Event-based Prospective Memory in School-aged Children." *Journal of Experimental Child Psychology* 166: 17–33. https://doi.org/10.1016/j.jecp.2017.08.002.

Dai, H., B. J. Dietvorst, B. Tuckfield, K. L. Milkman, and M. E. Schweitzer. 2018. "Quitting When the Going Gets Tough: A Downside of High Performance Expectations." *Academy of Management Journal* 61: 1667–1691.

Davis, J., E. Cullen, and T. Suddendorf. 2015. "Understanding Deliberate Practice in Preschool Aged Children." *Quarterly Journal of Experimental Psychology* 69: 1–43. https://doi.org/10.1080/17470218.2015.1082140.

de Hevia, M. D., and E. S. Spelke. 2010. "Number-Space Mapping in Human Infants." *Psychological Science* 21: 653–660. https://doi.org/10.1177/0956797610366091.

Dweck, C. S. 2006. *Mindset: The New Psychology of Success*. Random House Publishing Group.

Dweck, C. S., and D. S. Yeager. 2019. "Mindsets: A View From Two Eras." *Perspectives on Psychological Science* 14: 481–496. https://doi.org/10.1177/1745691618804166.

Evans, N. J., S. D. Brown, D. J. K. Mewhort, and A. Heathcote. 2018. "Refining the Law of Practice." *Psychological Review* 125: 592–605. https://doi.org/10.1037/rev0000105.

Finn, B., and J. Metcalfe. 2008. "Judgments of Learning Are Influenced by Memory for Past Test." *Journal of Memory and Language* 58: 19. https://doi.org/10.1016/j.jml.2007.03.006.

Flavell, J. H., A. G. Friedrichs, and J. D. Hoyt. 1970. "Developmental Changes in Memorization Processes." *Cognitive Psychology* 1: 324–340. https://doi.org/10.1016/0010-0285(70)90019-8.

Gaspelin, N., T. Margett-Jordan, and E. Ruthruff. 2015. "Susceptible to Distraction: Children Lack Top-Down Control Over Spatial Attention Capture." *Psychonomic Bulletin & Review* 22: 461–486. https://doi.org/10.3758/s13423-014-0708-0.

Gathercole, S. E. 1998. "The Development of Memory." *Journal of Child Psychology and Psychiatry* 39: 3–27. https://doi.org/10.1017/S0021963097001753.

Ghetti, S., E. Hembacher, and C. A. Coughlin. 2013. "Feeling Uncertain and Acting on It During the Preschool Years: A Metacognitive Approach." *Child Development Perspectives* 7: 160–165. https://doi.org/10.1111/cdep.12035.

Gonzales, C. R., A. Merculief, M. M. McClelland, and S. Ghetti. 2022. "The Development of Uncertainty Monitoring During Kindergarten: Change and Longitudinal Relations With Executive Function and Vocabulary in Children From Low-Income Backgrounds." *Child Development* 93: 524–539. https://doi.org/10.1111/cdev.13714.

Gottlieb, J., and P.-Y. Oudeyer. 2018. "Towards a Neuroscience of Active Sampling and Curiosity." *Nature Reviews Neuroscience* 19: 758–770. https://doi.org/10.1038/s41583-018-0078-0.

Goupil, L., M. Romand-Monnier, and S. Kouider. 2016. "Infants Ask for Help When They Know They Don't Know." *Proceedings of the National Academy of Sciences* 113: 3492–3496. https://doi.org/10.1073/pnas.1515129113.

Gweon, H., M. Asaba, and G. Bennett-Pierre. 2017. "Reverse-Engineering the Process: Adults' and Preschoolers' Ability to Infer the Difficulty of Novel Tasks." Proceedings of the 39th Annual Conference of the Cognitive Science Society.

Haimovitz, K., and C. S. Dweck. 2017. "The Origins of Children's Growth and Fixed Mindsets: New Research and a New Proposal." *Child Development* 88: 1849–1859. https://doi.org/10.1111/cdev.12955.

Heathcote, A., S. Brown, and D. J. K. Mewhort. 2000. "The Power Law Repealed: The Case for an Exponential Law of Practice." *Psychonomic Bulletin & Review* 7: 185–207. https://doi.org/10.3758/BF03212979.

Hennefield, L., and L. Markson. 2022. "The Development of Optimistic Expectations in Young Children." *Cognitive Development* 63: 101201. https://doi.org/10.1016/j.cogdev.2022.101201.

Henrich, J., S. J. Heine, and A. Norenzayan. 2010. "Most People Are Not WEIRD." *Nature* 466: 29a. https://doi.org/10.1038/466029a.

Hiscock, M., M. Kinsbourne, M. Samuels, and A. E. Krause. 1985. "Effects of Speaking Upon the Rate and Variability of Concurrent Finger Tapping in Children." *Journal of Experimental Child Psychology* 40: 486–500. https://doi.org/10.1016/0022-0965(85)90079-7.

Hunt, E. 1978. "Mechanics of Verbal Ability." *Psychological Review* 85, no. 2: 109–130. https://doi.org/10.1037/0033-295X.85.2.109.

Huttenlocher, J., N. C. Jordan, and S. C. Levine. 1994. "A Mental Model for Early Arithmetic." *Journal of Experimental Psychology: General* 123: 284–296. https://doi.org/10.1037/0096-3445.123.3.284.

Kominsky, J. F., T. Gerstenberg, M. Pelz, et al. 2021. "The Trajectory of Counterfactual Simulation in Development." *Developmental Psychology* 57: 253–268. https://doi.org/10.1037/dev0001140.

Krakauer, J. W., A. M. Hadjiosif, J. Xu, A. L. Wong, and A. M. Haith. 2019. "Motor Learning." In *Comprehensive Physiology*, edited by R. Terjung, 1st ed., 613–663. Wiley. https://doi.org/10.1002/cphy.c170043.

Lapidow, E., I. Killeen, and C. M. Walker. 2022. "Learning to Recognize Uncertainty vs. Recognizing Uncertainty to Learn: Confidence Judgments and Exploration Decisions in Preschoolers." *Developmental Science* 25, no. 2: e13178. https://doi.org/10.1111/desc.13178.

Leahy, B., M. Huemer, M. Steele, S. Alderete, and S. Carey. 2022. "Minimal Representations of Possibility at Age 3." *Proceedings of the National Academy of Sciences* 119: e2207499119. https://doi.org/10.1073/pnas.2207499119.

Leonard, J., and J. A. Sommerville. 2024. "A Unified Account of Why Optimism Declines in Childhood." *Nature Reviews Psychology* 4: 1–14. https://doi.org/10.1038/s44159-024-00384-z.

Leonard, J. A., S. R. Cordrey, H. Z. Liu, and A. P. Mackey. 2023. "Young Children Calibrate Effort Based on the Trajectory of Their Performance." *Developmental Psychology* 59: 609–619. https://doi.org/10.1037/dev0001467.

Lockhart, K. L., B. Chang, and T. Story. 2002. "Young Children's Beliefs About the Stability of Traits: Protective Optimism?" *Child Development* 73: 1408–1430. https://doi.org/10.1111/1467-8624.00480.

Lockhart, K. L., M. K. Goddu, and F. C. Keil. 2021. "How Much Can You Learn in One Year? How Content, Pedagogical Resources, and Learner's Age Influence Beliefs About Knowledge Acquisition." *Cognitive Development* 60: 101115. https://doi.org/10.1016/j.cogdev.2021.101115.

Luft, A. R., and M. M. Buitrago. 2005. "Stages of Motor Skill Learning." *Molecular Neurobiology* 32: 205–216. https://doi.org/10.1385/MN:32:3:205.

Mazur, J. E., and R. Hastie. 1978. "Learning as Accumulation: A Reexamination of the Learning Curve." *Psychological Bulletin* 85: 19.

Miller, P. H., W. L. Seier, J. S. Probert, and P. A. Aloise. 1991. "Age Differences in the Capacity Demands of a Strategy Among Spontaneously Strategic Children." *Journal of Experimental Child Psychology* 52: 149–165. https://doi.org/10.1016/0022-0965(91)90057-Y.

Muenks, K., A. Wigfield, and J. S. Eccles. 2018. "I Can Do This! The Development and Calibration of Children's Expectations for Success and Competence Beliefs." *Developmental Review* 48: 24–39. https://doi.org/10.1016/j.dr.2018.04.001.

Muradoglu, M., T. Porter, K. Trzesniewski, and A. Cimpian. 2024. "A Growth Mindset Scale for Young Children (GM-C): Development and Validation Among Children From the United States and South Africa." *PLoS One* 19: e0311205. https://doi.org/10.1371/journal.pone.0311205.

Newcombe, N. S. 2014. "The Origins and Development of Magnitude Estimation." *Ecological Psychology* 26: 147–157. https://doi.org/10.1080/10407413.2014.875333.

Nyhout, A., and P. A. Ganea. 2019. "Mature Counterfactual Reasoning in 4-and 5-Year-Olds." *Cognition* 183: 57–66. https://doi.org/10.1016/j.cognition.2018.10.027.

Rafetseder, E., M. Schwitalla, and J. Perner. 2013. "Counterfactual Reasoning: From Childhood to Adulthood." *Journal of Experimental Child Psychology* 114: 389–404. https://doi.org/10.1016/j.jecp.2012.10.010.

Roebers, C. M. 2017. "Executive Function and Metacognition: Towards a Unifying Framework of Cognitive Self-Regulation." *Developmental Review* 45: 31–51. https://doi.org/10.1016/j.dr.2017.04.001.

Rossi, C., C. W. Chau, K. A. Leech, M. A. Statton, A. J. Gonzalez, and A. J. Bastian. 2019. "The Capacity to Learn New Motor and Perceptual Calibrations Develops Concurrently in Childhood." *Scientific Reports* 9: 9322. https://doi.org/10.1038/s41598-019-45074-6.

Rule, J., M. K. Goddu, J. Chu, et al. 2023. "Fun Isn't Easy: Children Selectively Manipulate Task Difficulty When "Playing for Fun" vs. "Playing to Win"." https://doi.org/10.31234/osf.io/q7wh4.

Schäfer, S. 2005. "Concurrent Cognitive and Sensorimotor Performance: A Comparison of Children and Young Adults." https://doi.org/10.17169/refubium-13618.

Schneider, W. 1998. "Performance Prediction in Young Children: Effects of Skill, Metacognition and Wishful Thinking." *Developmental Science* 1: 291–297. https://doi.org/10.1111/1467-7687.00044.

Schneider, W. 2008. "The Development of Metacognitive Knowledge in Children and Adolescents: Major Trends and Implications for Education." *Mind, Brain, and Education* 2: 114–121. https://doi.org/10.1111/j.1751-228X.2008.00041.x.

Schulz, L. E., and E. B. Bonawitz. 2007. "Serious Fun: Preschoolers Engage in More Exploratory Play When Evidence Is Confounded." *Developmental Psychology* 43: 1045–1050. https://doi.org/10.1037/0012-1649.43.4.1045.

Serko, D., J. Leonard, and A. Ruggeri. 2024. "Children Strategically Decide What to Practice." OSF. https://doi.org/10.31234/osf.io/kzjtn.

Setoh, P., R. M. Scott, and R. Baillargeon. 2016. "Two-And-A-Half-Year-Olds Succeed at a Traditional False-Belief Task With Reduced Processing Demands." *Proceedings of the National Academy of Sciences* 113: 13360–13365. https://doi.org/10.1073/pnas.1609203113.

Sobel, D. M., and S. M. Letourneau. 2015. "Children's Developing Understanding of What and How They Learn." *Journal of Experimental Child Psychology* 132: 221–229. https://doi.org/10.1016/j.jecp.2015.01.004.

Solum, M., H. Lorås, and A. V. Pedersen. 2020. "A Golden Age for Motor Skill Learning? Learning of an Unfamiliar Motor Task in 10-Year-Olds, Young Adults, and Adults, When Starting From Similar Baselines." *Frontiers in Psychology* 11: 538. https://doi.org/10.3389/fpsyg.2020.00538.

Son, L. K., and R. Sethi. 2006. "Metacognitive Control and Optimal Learning." *Cognitive Science* 30: 759–774. https://doi.org/10.1207/s15516709cog0000_74.

Ten, A., P. Kaushik, P.-Y. Oudeyer, and J. Gottlieb. 2021. "Humans Monitor Learning Progress in Curiosity-Driven Exploration." *Nature Communications* 12: 5972. https://doi.org/10.1038/s41467-021-26196-w.

Turoman, N., R. I. Tivadar, C. Retsa, A. M. Maillard, G. Scerif, and P. J. Matusz. 2021. "The Development of Attentional Control Mechanisms in Multisensory Environments." *Developmental Cognitive Neuroscience* 48: 100930. https://doi.org/10.1016/j.dcn.2021.100930.

Vasilyeva, M., and S. F. Lourenco. 2012. "Development of Spatial Cognition." *Wiley Interdisciplinary Reviews: Cognitive Science* 3: 349–362. https://doi.org/10.1002/wcs.1171.

Xia, M., A. M. G. Poorthuis, and S. Thomaes. 2023. "Why Do Young Children Overestimate Their Task Performance? A Cross-Cultural Experiment." *Journal of Experimental Child Psychology* 226: 105551. https://doi.org/10.1016/j.jecp.2022.105551.

Xia, M., A. M. G. Poorthuis, and S. Thomaes. 2024. "Children's Overestimation of Performance Across Age, Task, and Historical Time: A Meta-Analysis." *Child Development* 95: 1001–1022. https://doi.org/10.1111/cdev.14042.

Yeager, D. S., P. Hanselman, G. M. Walton, et al. 2019. "A National Experiment Reveals Where a Growth Mindset Improves Achievement." *Nature* 573, no. 7774: 364–369. https://doi.org/10.1038/s41586-019-1466-y.

Yussen, S. R., and V. M. Levy. 1975. "Developmental Changes in Predicting One's Own Span of Short-Term Memory." *Journal of Experimental Child Psychology* 19: 502–508. https://doi.org/10.1016/0022-0965(75)90079-X.

Zhang, X., S. D. McDougle, and J. A. Leonard. 2025. "People Accurately Predict the Shape of Skill Learning Curves." *Cognition 258: 106083.* https://doi.org/10.1016/j.cognition.2025.106083.

Supporting Information

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