

# Developmental Psychology

## Young Children Calibrate Effort Based on the Trajectory of Their Performance

Julia A. Leonard, Skyler R. Cordrey, Hunter Z. Liu, and Allyson P. Mackey

Online First Publication, September 29, 2022. <http://dx.doi.org/10.1037/dev0001467>

### CITATION

Leonard, J. A., Cordrey, S. R., Liu, H. Z., & Mackey, A. P. (2022, September 29). Young Children Calibrate Effort Based on the Trajectory of Their Performance. *Developmental Psychology*. Advance online publication. <http://dx.doi.org/10.1037/dev0001467>

# Young Children Calibrate Effort Based on the Trajectory of Their Performance

Julia A. Leonard<sup>1,2</sup>, Skyler R. Cordrey<sup>1</sup>, Hunter Z. Liu<sup>1</sup>, and Allyson P. Mackey<sup>1</sup>

<sup>1</sup>Department of Psychology, School of Arts and Sciences, University of Pennsylvania

<sup>2</sup>Department of Psychology, Yale University

Learning requires effort, but children cannot try hard at everything. Here, we evaluated whether children use their improvement over time to decide whether to stick with a challenge. To eliminate the effect of individual differences in ability or prior knowledge, we created a novel paradigm that allowed us to surreptitiously control children's performance. Across three preregistered experiments ( $N = 319$ , ages 4 to 6 in the United States), we found that children who were given evidence that their performance was improving were more likely to persist on a challenging task than children who were given evidence that their performance was constant, even when final performance was matched. This effect was robust to differing reward contingencies, across in-person and online testing contexts, and was driven by the demotivating effect of constant performance. Our results suggest that young children will be more persistent if they are guided away from too-difficult tasks and toward opportunities that enable steady growth.

### Public Significance Statement

How do children decide which challenges are worth pursuing? We show that young children pay attention to their rate of past performance: 4- to 6-year-olds are more likely to persist with a challenge if their past performance increases over time versus stays constant, even when their final score is the same. Our results suggest that adults can help foster young children's persistence by guiding children toward achievable tasks just outside their comfort zone and highlighting their growth.


**Keywords:** early childhood, motivation, persistence, rational learning

**Supplemental materials:** <https://doi.org/10.1037/dev0001467.supp>

The road to mastery is riddled with setbacks. Even though failure is aversive, sometimes we decide to keep going. Understanding how children make decisions about effort is important, because persistence sets the stage for a broad range of positive life outcomes including academic success and strong interpersonal relationships (Banerjee & Tamis-LeMonda, 2007; Eskreis-Winkler et al., 2014; Martin et al., 2013; Messer et al., 1986; Mischel et al., 1989). Once we understand how children decide when to persist

and when to give up, we can use this knowledge to develop strategies that help people of all ages stick with their goals, build effective online learning platforms, and even design intelligent agents that learn tractable problems.

One key factor that may impact persistence in young children is *learning progress*: Children may feel more motivated to pursue tasks when they are improving, compared with when their performance is flat. This strategy is rational under theories of optimal

Julia A. Leonard  <https://orcid.org/0000-0001-8099-2721>

We thank A. Nerenberg, G. Bizzell-Hatcher, J. Sandler, and A. Rubio for help with data collection and coding, A. Schechter for help with building stimuli, and C. Recto for help with figure graphics. We thank R. Magid, L. Schulz, N. Vélez, and S. Sharp for helpful conversations and insightful comments. We are grateful for the Please Touch Museum and the children and families who participated in this research. This research was supported by a Jacobs Foundation Early Career Research Fellowship to Allyson P. Mackey, National Institute on Drug Abuse (1R34DA050297-010) to Allyson P. Mackey, and a University of Pennsylvania mindCORE Postdoctoral Fellowship to Julia A. Leonard. The authors declare no competing interests.

All experiments were preregistered. All reported data, analysis code,

and materials are available on the Open Science Framework ([https://osf.io/kv9jy/?view\\_only=3e5a657c65884be0982c437082e1e7c3](https://osf.io/kv9jy/?view_only=3e5a657c65884be0982c437082e1e7c3)). The Institutional Review Board at the University of Pennsylvania approved all experiments.

Julia A. Leonard contributed equally to conceptualization, data curation, writing—original draft, and writing—review and editing. Skyler R. Cordrey contributed equally to data curation. Hunter Z. Liu served as lead for data curation and served in a supporting role for software. Allyson P. Mackey served as lead for software and supervision and contributed equally to conceptualization, writing—original draft, and writing—review and editing.

Correspondence concerning this article should be addressed to Julia A. Leonard, Department of Psychology, Yale University, New Haven, CT 06520, United States. Email: [julia.leonard@yale.edu](mailto:julia.leonard@yale.edu)

learning (Metcalf & Jacobs, 2010; Metcalf & Kornell, 2005; Son & Sethi, 2006, 2010) and has already been implemented in some curiosity-driven intelligent machines and algorithms (Kim et al., 2020; Oudeyer et al., 2007; Wang et al., 2020). Imagine one child playing a game in which they get progressively closer to the goal each time they play. Imagine another child playing the same game, but they stay similarly close to reaching the goal each time they play. Which child will keep playing the game? Theories of optimal learning would predict that the child who is getting better over time should keep playing the game, as they are more likely to reach their goal on future attempts. Seeing oneself get better may also lead to feelings of pride and a sense of progress (Gürel et al., 2020), which may further motivate one to keep going. However, to date, most evidence in support of this prediction is correlational and centers on adults (e.g., Metcalf & Kornell, 2005; Payne et al., 2007; Ten et al., 2021). Thus, two fundamental questions remain: Does learning progress *causally* impact decisions about effort? Do *children* have the capacity to strategically allocate effort based on past performance?

There is evidence that even very young children can rationally detect when their efforts will pay off. For example, 13- to 18-month-old infants try harder on tasks after watching an adult effortfully versus effortlessly succeed, but rationally opt *not* to try if they see that an adult's efforts do not pay off (Leonard et al., 2017; Lucca et al., 2020). Infants also weigh rewards: 12- to 14-month-olds selectively crawl toward a jar with the higher proportion of their favored lollipops (Denison & Xu, 2010). However, in these studies, children reasoned about external evidence, not their prior behavior, when deciding how to deploy their effort.

It is less clear whether young children can accurately reason about their past performance. On the one hand, even infants demonstrate rudimentary metacognitive abilities and ask for help when they are uncertain (Goupil et al., 2016). Later in development, 4- to 5-year-olds demonstrate metacognitive awareness by reporting more confidence for items that they answer correctly versus incorrectly (Hembacher & Ghetti, 2014) and by proactively selecting evidence that will be easier for them to discriminate (Siegel et al., 2021). However, some aspects of metacognition may be later developing (Metcalf & Finn, 2013; Niebaum & Munakata, 2020). For example, after doing poorly on a memory test, 4- to 5-year-old children do not allocate more study time when presented with the same task again (Flavell et al., 1970). Older children (ages 5–7), but not younger children (ages 3–4), chose to be tested on easier numerical discrimination items, suggesting that aspects of metacognitive reasoning improve with age (Baer & Odic, 2019). In short, it is unclear whether young children are capable of using the trajectory of their past performance to predict their future performance and calibrate their effort accordingly.

### The Current Study

Here, we tested whether 4- to 6-year-old children use the trajectory of their past performance to inform decisions about whether to persist with a challenge. To eliminate the confounds of individual differences in skill or prior experience, we created a novel, intuitive game in which we could surreptitiously control children's performance. Across three preregistered experiments, two in-person and one online, we randomly assigned children to a condition in which they improved at a game, or a condition in which they stayed about the

same, with matched final performance. After four trials, we asked whether children wanted to keep playing with the same challenging version of the game, or switch to playing with an easier version. In Experiments 1 and 2, we tested how reward contingencies interact with information about past performance to inform children's decisions. In Experiment 3, we examined how children's performance expectations changed after receiving performance feedback and whether our results generalized across contexts. We hypothesized that children would be more likely to choose to stick with the challenging version of the task if they got better over time versus if their performance stayed flat. In sum, we sought to establish clear, causal evidence that performance curves impact young children's decisions to persist with challenges.

## Experiment 1

### Method

The Institutional Review Board at the University of Pennsylvania approved all experiments. Parents provided informed, written consent for their children's participation in all experiments. In Experiments 1 and 2, participants were recruited at an urban children's museum and tested individually in quiet testing rooms off the museum floor. The preregistration for Experiment 1 is here: [https://osf.io/vmu6e/?view\\_only=dc99339ec95d456a8a54e0972e6ee969](https://osf.io/vmu6e/?view_only=dc99339ec95d456a8a54e0972e6ee969); see minor deviations from preregistration in the online supplemental materials).

### Participants

Seventy-three 4- to 6-year-old children were recruited for the study, but only 66 were included in the data analysis (mean age: 63.77 months; range: 48–83 months<sup>1</sup>) because of the preregistered exclusion criteria of missing video recording ( $n = 2$ ), parental interference ( $n = 2$ ), or parent reporting diagnosed Autism Spectrum Disorder (ASD;  $n = 3$ ; note, owing to an oversight, ASD diagnosis was not preregistered exclusion criteria. Main results do not change with ASD participants included). The racial and ethnic makeup of the final sample was as follows: 41% White, 24% Hispanic/Latino, 14% Black, 14% Other, 12% Asian, and 8% Multiracial (missing race data from eight participants and ethnicity data from 11 participants). Parental education ranged from 12 to 20 years ( $M = 16.47$ ,  $SD = 2.51$ ; missing data from 6 participants; 68% holding four-year college degree). We preregistered running 33 subjects per condition from a power analysis anticipating a medium-large effect ( $V = .4$ ) and a power of .9. Children were randomly assigned to one of two conditions: Increasing (16 F) or Constant (22 F); age was matched across conditions,  $t(64) = -.49$ ,  $p = .63$ . All data were double scored by researchers blind to condition for all dependent variables and exclusion criteria with 100% agreement.

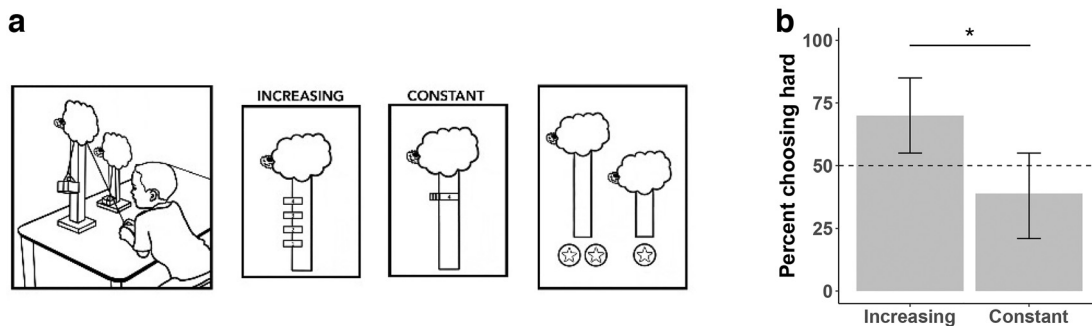
### Procedure

Children were introduced to two trees (one 17.5-in. tree and one 11.5-in. tree; see Figure 1a) and told that an egg had fallen out of a nest and that their goal was to put the egg back in a nest (inspired by a classic study by Stipek et al., 1984). The experimenter told

<sup>1</sup> Age was coded in months throughout.

**Figure 1**

Children Chose to Play With the Hard Tree More Often in the Increasing Condition Than in the Constant Condition in Experiment 1, When Rewards Were Higher for the Hard Versus Easy Tree



*Note.* (a) In Experiment 1, children tried to get an “egg” (a metal ball) back to a nest in a wooden tree with a pulley device. Unbeknownst to the child, the egg stayed on the platform via a hidden magnet in the platform that the experimenter could turn on and off via remote control. Children were randomly assigned to the Increasing condition, where the ball fell off at increasing higher intervals, or the Constant condition, where the ball fell off around the same point. Children could receive two stickers for getting the egg back to the nest in the harder tree and one sticker for getting the egg back to the nest in the easier tree. (b) Children were more likely to stick with a challenge in the Increasing condition than the Constant condition. Error bars represent bootstrapped 95% confidence intervals.

\*  $p < .05$ .

children that they would receive two stickers for getting the egg to the nest in the tall tree or one sticker for getting the egg to the nest in the small tree (unequal reward). The experimenter said that they would start by playing with the tall tree. Unbeknownst to the child, the egg (a metal ball) stayed on the platform via a hidden electromagnet in the platform which the experimenter could turn on and off with a remote control. While children pulled up the platform, the ball wobbled, making it appear as a conceivably difficult balancing task, and fell off at prespecified heights. Children played with the tall tree for four consecutive trials. In the Increasing condition, the experimenter surreptitiously made the ball fall off the platform at 8", 10", 12", 14" sequentially during each respective trial. In the Constant condition, the ball fell off around 14" on each trial. After the ball fell off on each trial, the experimenter said, “Oops, it fell off! But good job!,” then marked children’s progress by placing a marker with the trial number on the Velcro trunk at the height where the ball fell off and prompted the child to try again. After the fourth trial, the experimenter reviewed the child’s performance by pointing to how high they got the egg on each trial. Finally, the experimenter asked the critical question, “You’re so close to getting to the top of this tree! Do you want to try this tree again (pointing to the tall tree) or try the other tree (pointing to the small tree)?”

Importantly, this paradigm capitalizes on children’s intuitive understanding of relative distance (Boyer & Levine, 2015; de Hevia & Spelke, 2010; Huttenlocher et al., 1994; Newcombe, 2014). To check whether children thought the game was rigged, we asked children to explain their final tree choice. In general, children believed the manipulation: No child said that they chose to switch to the small tree because the tall tree was broken or rigged. This was true across all three experiments. We also asked children whether they thought they got better or stayed the same at the game over time to check children’s interpretation of the manipulation (scored as binary: 1 = response consistent with condition manipulation, 0 = response inconsistent with condition manipulation). Finally, the experimenter

ensured that the child reached the nest on their chosen tree. All sessions were video recorded, and data were double scored by researchers blind to condition for all dependent variables and exclusion criteria with 100% agreement.

## Results

The critical comparison was between children’s choices to stick with a challenge in the Increasing condition versus the Constant condition. Children in both conditions had failed four times and ended up in the same place, but the conditions differed in slope of progress. As predicted, a chi-square test revealed that children in the Increasing condition chose to continue playing with the hard tree more often than children in the Constant condition,  $\chi^2(1, n = 66) = 4.95, p = .03, V = .30$ ; Figure 1b). Children in the Increasing condition chose to stick with a challenge significantly *above* chance (binomial test against chance (50%): 70% hard, 95% CI [55%, 85%],  $p = .04$ ). Children’s choices in the Constant condition did not differ from chance (binomial test against chance (50%): 39% hard, 95% CI [21%, 55%],  $p = .30$ ).

Children understood the condition manipulation. A chi-square test revealed that children in the Increasing condition were more likely to say they got better at the game over time than children in the Constant condition,  $\chi^2(1, n = 59) = 9.4, p = .002, V = .43$ ; no data from seven participants). Children in the Constant condition were above chance at endorsing that they stayed the same at the game (binomial test against chance (50%): 79% endorse “staying the same,”  $p = .004$ ), whereas children in the Increasing condition did not choose “getting better” at the game above chance (binomial test against chance (50%): 65% endorse “getting better,”  $p = .15$ ). A linear model predicting whether participants got the manipulation check correct revealed no effect of age ( $b = .008, p = .19$ ).

Next, we explored age effects on tree choice. In a logistic regression predicting tree choice with age and condition, both age and condition were significant (age:  $b = .06, p = .02$ ; condition:

$b = 1.31, p = .02$ ), revealing that older children and children in the Increasing condition were more likely to choose the hard tree. A model that included an age by condition interaction on tree choice revealed no significant interaction ( $b = -.03, p = .57$ ).

## Discussion

Results from our first experiment show that children are sensitive to their rate of improvement: Children who got closer to reaching a goal over time were more likely to stick with a challenging task than children who were consistently close to reaching a goal over time, even when the distance to the goal on the final trial was matched. Older children were also more likely to stick with a challenge across conditions. Children chose to stick with a challenge above chance only when their performance was improving over time. However, reaching the more challenging goal was associated with a higher reward (two stickers vs. one), so it is unclear whether increasing performance in the absence of higher payoff will motivate children to stick with a challenge. To answer this question, and replicate our condition differences, we ran a second experiment where we offered the same reward for succeeding on both the tall and small tree. Additionally, we added a Baseline no manipulation condition to further interrogate the directionality of our condition differences.

## Experiment 2

In this experiment (preregistration: [https://osf.io/dkyfe/?view\\_only=9806a0e78c8c45618a6e5ae723643057](https://osf.io/dkyfe/?view_only=9806a0e78c8c45618a6e5ae723643057)) children were randomly assigned to the Increasing, Constant, or Baseline (no prior experience) condition and told that they could get one sticker for getting the egg back to the nest in the tall tree or small tree (matched reward). We predicted that we would replicate our condition difference from Experiment 1, with children in the Increasing condition being more likely to stick with a challenge than children in the Constant condition. However, we were unsure about whether children in the Increasing condition would stick with a challenge more than children in the Baseline condition or chance (50%). On

the one hand, children in the Increasing condition might find increasing performance motivating in and of itself. On the other hand, children may integrate information about their past performance with potential rewards and thus be split between sticking with or abandoning a challenge (their probability of success and rewards are essentially matched across trees). However, we predicted that children in the Constant condition should now be more likely than chance to choose the small tree because their chance of success is higher on this tree and rewards are matched across options.

## Method

### Participants

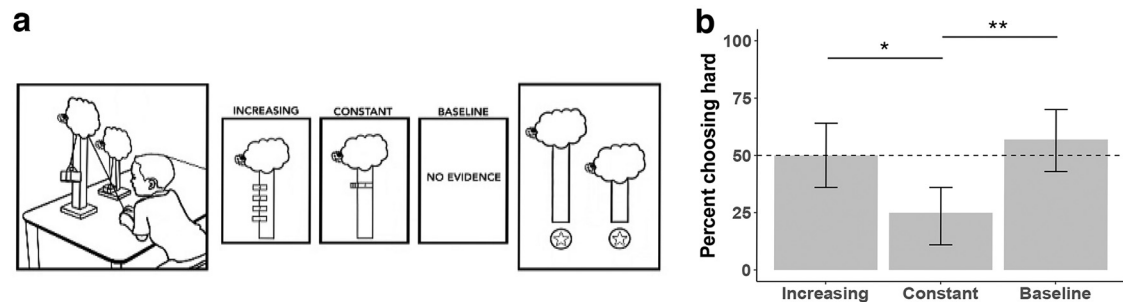
In Experiment 2, 140 four- to 6-year-old children were recruited, but only 132 were included in the data analysis (mean age: 65.34 months; range: 48–83 months) owing to the preregistered exclusion criteria of meaningful distraction at the museum ( $n = 1$ ), watching a friend do the experiment ( $n = 1$ ), experimental error ( $n = 2$ ), or parent reporting diagnosed ASD ( $n = 4$ ). The racial and ethnic makeup of the final sample was as follows: 59% White, 14% Black, 11% Asian, 11% Hispanic/Latino, 8% Multiracial, and 5% Other (missing race data from four participants and ethnicity data from six participants). Parental education ranged from 12 to 20 years ( $M = 16.23, SD = 2.62$ ; missing data from two participants; 71% holding four-year college degree). Children were randomly assigned to one of three conditions: Increasing (24 F), Constant (29 F), and Baseline (26 F); age was matched across conditions (age:  $F[2, 129] = .091, p = .91$ ). Based on an expected medium effect size for comparing two conditions at a time ( $V = .3$ ), we preregistered running 44 subjects per condition with power of .8.

### Procedure

The procedure for the Increasing and Constant conditions was the same as in Experiment 1 except that children were offered one sticker for getting the egg to the nest in either tree (Figure 2a) and, as an additional measure of children's perception of their task

**Figure 2**

*Children Chose to Play With the Hard Tree More Often in the Increasing Condition Than in the Constant Condition in Experiment 2, When Rewards Were Equal for Both Trees*



*Note.* (a) Experiment 2 was the same as Experiment 1 except children could receive one sticker for getting the egg back to the nest in the harder tree or easier tree and there was a Baseline condition. (b) Children were more likely to play with the hard tree in the Increasing condition than in the Constant condition. There was no difference between children's choices in the Increasing and Baseline condition, however children in the Constant condition chose to play with the easy tree more than children in the Baseline condition. Error bars represent bootstrapped 95% confidence intervals.

\*  $p < .05$ . \*\*  $p < .01$ .

performance, children were asked to make an ability judgment prior to the critical task choice question. For the ability judgment, the experimenter took out a piece of paper with a big thumbs up and down and pointed to the corresponding symbols when asking children whether they thought they were good or bad at the game (order counterbalanced). If children said “good,” they were shown a big and small thumbs up and asked if they thought they were “somewhat good” (small thumbs up) or “really good” (big thumbs up) at the game. Similarly, if children said “bad,” they were asked if they were “somewhat bad” (small thumbs down) or “really bad” (big thumbs down) at the game. In the Baseline condition, children did not play with the tree at all before being asked “You get to choose—which tree do you want to play with?” Children were not explicitly told that one tree was harder than the other in the Baseline condition but prior work suggests that preschool age children are able to infer task difficulty from physical features (Bennett-Pierre et al., 2018; Gweon et al., 2017).

## Results

As in Experiment 1, a chi-square test revealed that children in the Increasing condition chose to continue playing with the hard tree more than children in the Constant condition,  $\chi^2(1, n = 88) = 4.85, p = .03$ , Cramer’s  $V = .26$  (Figure 2b). Following preregistered analyses to explore the directionality of effects, we ran chi-square tests comparing children’s choices in each condition to Baseline and to chance performance using binomial tests. Children’s choices in the Increasing condition did not differ from those in the Baseline condition,  $\chi^2(1, n = 88) = .18, p = .67, V = .07$ . In both the Increasing and Baseline conditions, children’s choices did not differ from chance (Baseline: binomial test against chance [50%]: 57% hard, 95% CI [43%, 70%],  $p = .45$ , Increasing: 50% hard, 95% CI [36%, 64%],  $p = 1.0$ ). However, children in the Constant condition chose to continue playing with the harder tree less than children in the Baseline condition,  $\chi^2(1, n = 88) = 7.94, p = .005, V = .32$ , and significantly below chance (binomial test against chance (50%): 25% hard, 95% CI [11%, 36%],  $p = .001$ ).

Most children understood the condition manipulation. A chi-square test revealed a trend for children in the Increasing condition to say that they got better at the game over time more than children in the Constant condition,  $\chi^2(1, n = 85) = 2.66, p = .10$  (no data from three participants). However, within condition, children did not choose the correct response above chance (binomial test against chance [50%]: 58% endorse “getting better” in Increasing  $p = .36$ ; 62% endorse “staying the same” in Constant  $p = .16$ ). A binomial regression revealed that older children were more likely to correctly match their performance (“getting better” vs. “staying the same”) to their condition ( $b = .05, p = .04$ ). Children’s subjective judgements of ability did not differ by condition in a chi-square,  $\chi^2(2, n = 123) = 2.05, p = .36$  (dropping the five responses to “somewhat bad” and 0 responses to “really bad” to fulfill requirements for a chi-square; missing data from four participants). A binomial regression controlling for condition showed that children’s ability assessments decreased with age—younger children were more likely to say that they were “really good” at the game than older children ( $b = -.05, p = .01$ ).

Next, we examined age effects on tree choice within and across conditions. A linear model predicting tree choice by condition and age revealed main effects of condition, but not age (Constant  $b =$

$-1.38, p = .003$ ; Increasing  $b = -.27, p = .53$ ; age:  $b = .01, p = .58$ ). A model that included an age by condition interaction on tree choice revealed a significant interaction of age for the Constant (versus Baseline) condition ( $b = .09, p = .04$ ) but not the Increasing (versus Baseline) condition ( $b = .08, p = .07$ ). Specifically, there was a positive nonsignificant association between age and choosing the hard tree in the Constant condition ( $b = .05, p = .15$ ) and a nonsignificant negative association between age and choosing the hard tree in the Baseline condition ( $b = -.04, p = .15$ ).

## Discussion

Results from Experiment 2 again show that children are sensitive to their past performance: Children whose performance improved were more likely to stick with a challenge than children whose performance stayed the same. However, unlike in Experiment 1, we did not find that children’s verbal self-assessments of performance or ability tracked with condition. Instead, children’s verbal assessments of ability tracked with age more than condition, in line with work showing that younger children are more optimistic about their ability than older children (Flavell et al., 1970; O’Leary & Sloutsky, 2019; Parsons & Ruble, 1977; Schneider, 1998; Stipek & Hoffman, 1980; Xia et al., 2021). Even so, children’s behavioral choices still differed by condition, showing that children are sensitive to condition effects in ways not captured in their verbal responses.

Results from Experiment 2 also suggest that children weigh their likelihood of reaching a goal with the associated reward. Only when rewards were higher for the harder goal (Experiment 1, but not 2) did children choose to stick with the harder goal above chance when their performance was improving. Children whose performance stayed flat were more likely to switch to an easier goal compared with children who had no initial experience of failure when reward were equal among outcomes. There was no difference in children’s choices between the Increasing condition and Baseline condition. Thus, the condition difference between the Increasing and Constant condition when rewards are equal appears to be driven by the demotivating impact of stagnant performance.

However, given the disconnect between children’s verbal endorsements of their performance and their task choice by condition, it is unclear how exactly children are interpreting their performance. It is also unclear whether these results hold across different testing contexts. In Experiment 3 we sought to answer these questions. Specifically, we tested whether children would be more likely to predict increases in their performance over time in the Increasing condition versus the Constant condition. To lower task demands and control for variability in young children’s verbal abilities, we collected nonverbal predictions of future task performance via pointing. We also tested whether we could replicate our main condition effects (Increasing vs. Constant) from Experiment 2 in an online, at-home context. We focused just on the Increasing and Constant conditions because our main hypotheses concern this condition difference.

## Experiment 3

In Experiment 3, we explored children’s trial-by-trial predictions of performance and tested whether we could replicate condition differences from Experiment 2 (equal reward) in a home context

(preregistration: [https://osf.io/6jrfh/?view\\_only=805551e99ab34b81a6bc8cbae56fb668](https://osf.io/6jrfh/?view_only=805551e99ab34b81a6bc8cbae56fb668)). To this end, we adapted our paradigm to an online touchscreen game that children could play at home with parental supervision. The paradigm was the same as Experiment 2 except for the following changes: (a) we added a prediction phase before each trial, (b) we removed the Baseline condition, (c) we offered gold tokens instead of stickers for getting the egg to the nest, and (d) we had the egg fall off at slightly different places in the constant condition to avoid the game appearing rigged or broken (see Figure 2c).

## Method

### Participants

Participants were recruited through Facebook advertisements and ChildrenHelpingScience.com. Of a total of 159 participants, we excluded 38 owing to preregistered exclusion criteria: no audio recording ( $n = 17$ ), game malfunction midplay ( $n = 12$ ), second time playing game ( $n = 1$ ), incorrect age ( $n = 2$ ), audio recording not from a child ( $n = 6$ ). The final sample consisted of 121 four- to 6-year-old children (mean age: 65.31 months; range: 48–83 months; 57% F Increasing; 44% F Constant), which is more than our preregistered sample of 88 participants. We found that online data collection was less amenable to precise sampling because only some children play the game after the links are sent, and thus we had to oversample to get equal distributions of age within and across conditions. To match the preregistered sample size of 88, we selected the first 15 four-year-olds, first 14 five-year-olds, and first 15 six-year-olds who played in each condition. This age distribution also matched the age distribution in Experiment 2. We report results from the full sample in the main text and the preregistered smaller sample in the online supplemental materials (main results do not differ by sample). The racial makeup of the 121 subjects was as follows: 79% white, 12% Multiracial, 7% Asian, 2% other, and 2% prefer not to share. Parental education ranged from 10 to 20 years ( $M = 16.78$ ,  $SD = 2.12$ ; missing data from one participant; 86% holding four-year college degree). Participants came from 33

states, with the majority from PA ( $n = 19$ ) and CA ( $n = 11$ ); see the online supplemental materials for more information).

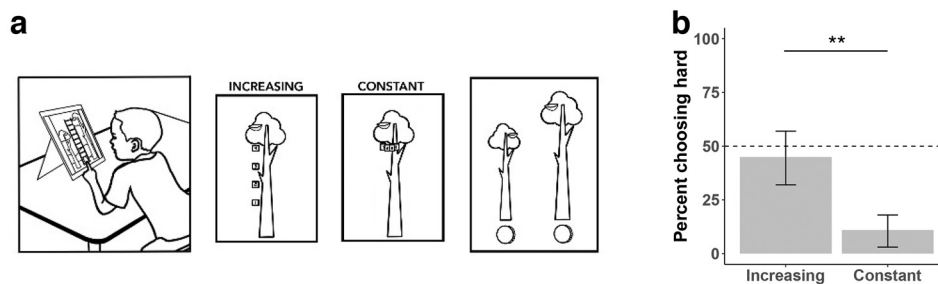
### Procedure

We created a virtual touch-screen version of the tree paradigm used in Experiments 1 and 2 using the JavaScript library React (<https://reactjs.org/>). We hosted the game on Heroku (<https://www.heroku.com/>) and used MongoDB (<https://www.mongodb.com/2>) for the database. To play, parents went to the website that hosted our game and entered their subject-specific ID and child birthdate. First, the game listed “Parent information” for the parent to read, specifying that they should stay with their child during the game because they would be asked to confirm their child’s choices at various points throughout the game. However, parents were instructed not to tell their child what to do or communicate with them by talking, using gestures, or cheering them on during the game. Throughout the game, children and parents were guided by audio recordings of the experimenter’s voice that cued certain images or buttons to appear. The next screen prompted parents to put the device on a stable surface and to make sure the Internet connection was good. The next screen asked children to confirm that a parent was around. Children were then asked to say their name and age out loud and then put their hands on their lap and pay careful attention. Children watched a quick video explaining the rules of the game (using similar language to the instructions used in Experiments 1 and 2). Children were told that they would get a point (shown as a golden token) for getting the egg back to either nest.

Children were randomly assigned to an Increasing condition or a Constant condition. Instead of using a pulley to bring the egg back to the nest, the online version of the task had children use their finger to lift the egg up on a platform mounted on a ladder to the nest (Figure 3a). Before children started playing the actual game, they could practice moving the platform that would hold the egg up and down the ladder with their finger. As in the in-person version of the Experiment, children had four trials to try to get the egg back to the nest. Before each trial, children were prompted to touch how high they thought they would get the egg on the tree. A star appeared at their guess and parents were asked to press “yes”

**Figure 3**

*Children Chose to Play With the Hard Tree More Often in the Increasing Condition Than in the Constant Condition in Experiment 3, an Online Version of the Experiment With Equal Rewards*



*Note.* (a) In Experiment 3, children played on a tablet device at home. Children were randomly assigned to an Increasing or Constant condition and made predictions about their performance before each trial. Rewards were matched across outcomes. (b) Children were more likely to stick with a challenge in the Increasing condition than the Constant condition. Error bars represent bootstrapped 95% confidence intervals.

\*\*  $p < .01$ .

to confirm their guess or “no” if the child touched the wrong spot (in which case, children were prompted to guess again). The star from their past guess disappeared once they made another guess. The position where the egg fell off was marked on the tree after each trial. In the Increasing condition, the egg fell off at 28%, 44%, 60%, and 76% to the top (of 100; same as the in-person intervals but scaled to 100) and in the Constant condition the egg fell off at 76%, 75%, 75%, and 76% at the top. After four trials, the “experimenter” went over the child’s past performance, using the same language as in Experiments 1 and 2. Finally, children were asked to touch which tree they wanted to play next. Parents confirmed the child’s choice by pressing “yes” or “no.”

Before children were able to play with their choice of tree for the final round, they were asked to say out loud why they picked that tree and to pick whether they thought that they were staying the same at the game (paired with a picture of a plant not growing) or getting better over time (paired with a picture of a plant growing; order counterbalanced). We decided not to ask about ability judgements owing to results from Experiment 2 showing that ability judgements were not sensitive to condition, but rather age. Finally, children could play with their desired tree, get the egg to the nest, and get a point. At the end of the game, parents could mark whether there was interference and had a chance to leave feedback. Any questions pertaining to interference or game play were followed up via email correspondence with parents.

To ensure that a child was actually playing the game, and not an adult, we recorded and analyzed children’s verbal responses to questions about the game. We had two coders mark whether they thought the audio recordings were from an adult or a child. Agreement was 100%. We only included children who had audio recordings.

## Results

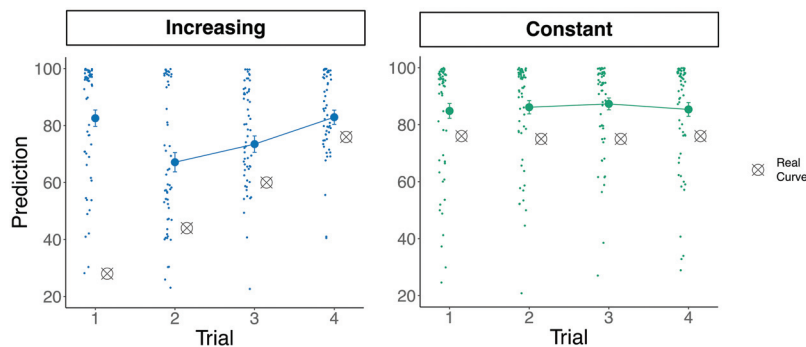
A chi-square test again revealed that children in the Increasing condition chose to continue playing with the hard tree more than children in the Constant condition,  $\chi^2(1, n = 121) = 15.21, p < .001, V = .37$  (Figure 3b). A binomial test (chance of 50%) showed

that children’s choices in the Increasing condition did not differ from chance (45% hard, 95% CI [32%, 57%],  $p = .52$ ), but children in the Constant condition chose to play with the hard tree significantly below chance (11% hard, 95% CI [3%, 18%],  $p < .001$ ). In a linear model predicting tree choice by age and condition, there was no effect of age (age:  $b = -.004, p = .27$ ; condition:  $b = .35, p < .01$ ). There was also no age by condition interaction in a linear model predicting tree choice ( $b = -.003, p = .69$ ). Thus, the online version of this paradigm robustly replicated condition differences from Experiment 2.

As in the in-person experiments, the paradigm was convincing to children. No child said that they chose to switch to the small tree because the tall tree was broken or rigged. A chi-square test showed that children in the Increasing condition were more likely to say they got better at the game over time than children in the Constant condition,  $\chi^2(1, n = 121) = 25.91, p < .001, V = .48$ . Binomial tests against chance (50%) showed that children in the Increasing condition endorsed getting better at the game above chance (88% endorse “getting better,”  $p < .001$ ), whereas children in the Constant condition did not endorse staying the same above chance (57% endorse “staying the same,”  $p = .31$ ). A linear model predicting whether children correctly answered the manipulation check question revealed no effect of age ( $b = .004, p = .34$ ).

Children made different predictions of their trial-by-trial performance based on the condition. Before children played the game, there was no difference between children’s predictions in the Increasing versus the Constant condition,  $t(119) = .59, p = .56$ , Increasing 1st prediction mean ( $SD$ ): 82.57 (22.43), Constant 1st prediction mean ( $SD$ ): 84.85 (20.22; Figure 4). Furthermore, in a linear model predicting children’s first predictions with age and condition, neither were significant (age:  $b = -.07, p = .72$ , condition:  $b = -2.09, p = .60$ ). However, a linear mixed-effects model across trials 2–4 revealed a main effect of condition ( $b = -29.23, p < .001$ ), and a condition by trial interaction ( $b = 8.98, p < .001$ ; Figure 4) on children’s performance predictions. Within the Increasing condition, a linear mixed-effects model revealed a positive main effect of trial on children’s performance predictions ( $b = 7.90, p < .001$ ). When age was added to the model, it was not

**Figure 4**  
Children’s Predictions of Performance in Experiment 3



*Note.* In the Increasing condition, children’s predictions changed after receiving performance feedback (“real curve”): Children predicted that their performance would improve across trials. In the Constant condition, children did not change their predictions of performance across trials (which can be expected when performance is constant). See the online article for the color version of this figure.



significant, and neither was an age by trial interaction. However, within the Constant condition, there was no main effect of trial on children's performance predictions ( $b = -.39$ ,  $p = .73$ ). Again, when age was added to the model, it was not significant, and neither was an age by trial interaction.

## Discussion

Even on a virtual task conducted in a more naturalistic home setting, children were more likely to stick with a challenge if they were increasingly close to reaching a goal over time than if they were consistently close to reaching a goal over time. Furthermore, we found evidence that children track their performance in accord with the condition manipulation: Children predicted that their performance would improve in the Increasing condition and stay the same in the Constant condition.

### General Discussion

We found that young children use the trajectory of their past performance to decide whether to persist with a challenge. In line with accounts of optimal learning (Metcalfe & Jacobs, 2010; Metcalfe & Kornell, 2005; Son & Sethi, 2006, 2010), children whose performance improved over time were more likely to stick with a challenge than children whose performance stayed the same. This was true even though the distance to the goal was matched on the final trial across conditions, and the overall final score was *higher* when children's performance was constant versus increasing. These findings held across in-person and online contexts and were driven by the demotivating effect of constant performance. In line with accounts of children as rational learners (Schulz, 2012; Tenenbaum et al., 2011), we showed that children are sensitive not only to their chance of getting a reward, but also the magnitude of that reward, when calibrating their effort. When rewards were equal across goals, children were at chance when deciding whether to stick with a challenge if their performance was improving. However, when rewards were higher for the harder goal, children chose to stick with a challenge above chance if they were getting better over time.

Many features of our experimental design may have been critical to children's successful reasoning. First, we developed a task that was appropriate and intuitive for preschoolers. By having children physically manipulate the task, which seemed tricky, yet possible, we led children to believe that they had agency over their performance. Indeed, children accurately predicted that their performance would improve in the Increasing condition and stay constant in the Constant condition. Furthermore, no child reported that the game was rigged or broken, and adding a jitter to the Constant condition did not change any main results. This is in line with prior literature showing that children have a bias toward attributing actions of others to individual, person-specific attributes rather than external situational factors (Seiver et al., 2013). We also used visual aids to lower memory demands and to help children track their own progress. Absent any of these features, children may not have been able to effectively monitor and act on their past performance.

Why do children stick with a challenge more when their performance is improving versus staying the same? Based on results from Experiment 3, we see that children notice and track their past

performance across both conditions and only predict that their future performance will improve in the Increasing condition. Thus, we believe that children's decisions to abandon a challenge in the Constant condition may reflect a rational inference that they are not likely to reach the goal on the next try. Similarly, children's decisions to stick with a challenge in the Increasing condition may reflect a rational inference that they are likely to improve on the next try. However, it could also be that children in the Increasing condition felt a sense of pride or inferred that their ability is malleable (see Gürel et al., 2020) which may have fueled their desire to stick with a challenge beyond the inference that their performance was improving. Similarly, children in the Constant condition may have experienced more boredom, monotony, or predictability, which may have motivated them to switch to a different task (Baker et al., 2010; Chin et al., 2017; Kidd et al., 2012, 2014; Pekrun et al., 2010). Future work with this paradigm should include measures of children's affective state and sense of self to understand more precisely how children interpret performance trajectories.

Results from Experiment 2 suggest that the condition difference between Increasing and Constant performance is driven by the demotivating effects of stagnant performance: Children's challenge preferences were only statistically different from Baseline in the Constant condition. In both the Increasing and Baseline condition, children's choices were at chance between the harder and easier goals. This pattern of results in the Increasing condition could be seen as rational because children could get the same reward for both goals and have the same chance of success on both goals. However, it is possible that providing explicit rewards may have dampened any intrinsic motivation children may have felt from their task improvement in the Increasing condition (Blain & Sharot, 2021; Deci, 1971). Thus, future work is needed to test whether children choose to stick with a challenge above chance when their performance is increasing in the absence of external rewards for performance outcomes. Furthermore, future work should explore whether children's preference for a challenge after increasing performance differs from children's preference for a challenge with no prior experience when they are offered a larger reward for the more challenging option (as in Experiment 1).

Age effects differed across experiments. When rewards were higher for the harder tree than the easier tree (as in Experiment 1), older children were more likely to choose the hard tree across conditions. This result is notable considering that we controlled children's performance, so older children did not perform better than younger children and should not have felt more competent. However, we did not find an effect of age on children's choices when rewards were matched across both outcomes (Experiment 2 and 3). We also did not find an impact of age on children's predictions of future performance across conditions, suggesting that older children are not more sensitive to their past performance data than younger children. Furthermore, in contrast to prior literature (Flavell et al., 1970; O'Leary & Sloutsky, 2019; Parsons & Ruble, 1977; Schneider, 1998; Stipek & Hoffman, 1980; Xia et al., 2021), we did not find that younger children were more optimistic about their future performance than older children (though note that we did find a predicted "optimism" age effect of younger children endorsing being "really good" at the task more than older children across conditions in Experiment 2). Thus, one possible explanation for older children choosing the harder tree across conditions in Experiment 1, but not 2 or 3, may be that older children

are more sensitive to reward discrepancies than younger children. However, future work is needed to more thoroughly test how age relates to reward sensitivity and learning from prior performance.

Children's verbal endorsements of "getting better" or "staying the same" were not consistent across Experiments. In Experiment 1 (unequal reward) and 3 (online, matched reward) children were more likely to say that they "got better" after four trials of playing the game in the Increasing condition versus the Constant condition. However, in Experiment 1 this effect was driven by children endorsing "staying the same" above chance in the Constant condition while in Experiment 3 this effect was driven by children endorsing "getting better" in the Increasing condition. Furthermore, there was a weak, nonsignificant condition difference in endorsements of "getting better" or "staying the same" in Experiment 2. Why are children's responses to this question so inconsistent? One possibility is that 4- to 6-year-olds might not totally grasp what these expressions mean or how they map onto the current task. For example, some children in Experiments 1 and 2 said that they stayed the same each time in the Increasing condition because the ball always dropped off—thus children may have been mapping this phrase to the feature of the ball consistently dropping rather than *where* it dropped. Furthermore, some children in the Constant condition claimed that they got better which may have accurately reflected very small deviations in where the ball fell off or a general optimistic bias in young children. Given the ambiguities inherent in this question, it is hard to draw strong conclusions about children's interpretation of the task from their verbal answers. However, the nonverbal prediction data in Experiment 3 suggest that children, on average, are sensitive to the condition effects.

An open question is whether increasing performance is more motivating than constant performance on tasks in which progress is opaque or tracked over days, months, or years. This question may be most pressing in a classroom context, where children's skills are constantly being evaluated. In 8- to 16-year-olds, comparing one's current performance to one's past performance rather than other people's performance effectively boosts pride and a desire for self-improvement rather than superiority over peers (Gürel et al., 2020). Thus, focusing on one's past performance curve may serve as an important intervention to enhance motivation in educational contexts. Future work should explore this intervention approach and test whether its effects are domain-general or specific (e.g., if you focus on self-growth on reading, do you also try harder on writing?; see Doan et al., 2020). But perhaps the more pressing question is how to motivate those whose performance has stayed flat.

In the real world, true performance feedback may be both motivating and demotivating, depending on a child's skill. For example, many digital games provide explicit evaluative feedback in the format of total progress, levels achieved, or time needed to complete a goal. There is even research on what type of progress bar is most motivating (e.g., completion bar vs. character building; Siemens et al., 2015). Our results suggest that games that track progress *over time*, rather than total score, may be the most motivating when performance is progressing and can lead to greater reward. However, this same type of feedback may backfire when performance is stagnant. What counts as progress also varies by task and domain. If you want to learn a new language, progress may mean the number of new words you know, and if you are

trying to get better at running, progress may mean the distance you are able to run. Both of these types of progress are easy to track and plot. However, in contexts where progress is more subjective or opaque, other forms of feedback, such as social accountability or tracking "streaks" in participation, may be more motivating (Dailey et al., 2018; Renfree et al., 2016).

This study has a number of limitations. First, our sample included only participants from a Western, educated, industrialized, rich, and democratic (WEIRD) country, thus we cannot address how our findings might differ across social contexts. Second, we controlled children's performance and thus performance on this task might not reflect naturalistic challenge-seeking behaviors. Third, asynchronous remote data collection is noisier than in-person data collection, and there may have been interruptions or interferences that were not captured. Despite this, we still observed the predicted condition difference in our online Experiment 3. Fourth, we said "Oops, it fell off! But good job!" after the ball fell off, but it is possible that children may have interpreted this feedback differently in each condition. Future work should explore how verbal feedback impacts how children perceive their past performance. Fifth, we defined improving performance as getting closer to an explicit, external goal, but feedback on progress could have different effects on motivation depending on the task, the skill, or the context. Finally, we only tested two learning curves, but it would be interesting to see whether children are sensitive to a larger range of performance curves (see Leonard et al., 2020).

As children grow and develop, they need to make calculated decisions about when to persist with challenges and when to move on and spend their effort on something more worthwhile. Here we show that children make rational decisions about their effort allocation based on their performance trajectory: 4- to 6-year-olds are more likely to stick with a challenge if their performance is improving over time versus constant over time. Our work suggests that educators and caregivers can encourage children to persevere by steering children away from too-difficult tasks and toward opportunities in which children can improve over time.

## References

- Baer, C., & Odic, D. (2019). Certainty in numerical judgments develops independently of the approximate number system. *Cognitive Development, 52*, Article 100817. <https://doi.org/10.1016/j.cogdev.2019.100817>
- Baker, R. S. J., D'Mello, S. K., Rodrigo, M., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4), 223–241. <https://doi.org/10.1016/j.ijhcs.2009.12.003>
- Banerjee, P. N., & Tamis-LeMonda, C. S. (2007). Infants' persistence and mothers' teaching as predictors of toddlers' cognitive development. *Infant Behavior and Development, 30*(3), 479–491. <https://doi.org/10.1016/j.infbeh.2006.12.001>
- Bennett-Pierre, G., Asaba, M., & Gweon, H. (2018). Preschoolers consider expected task difficulty to decide what to do and whom to help. In C. Kalish, M. Rau, J. Zhu, & T. Rogers (Eds.), *Proceedings of the 40th annual conference of the Cognitive Science Society* (pp. 1359–1364). Cognitive Science Society.
- Blain, B., & Sharot, T. (2021). Intrinsic reward: Potential cognitive and neural mechanisms. *Current Opinion in Behavioral Sciences, 39*, 113–118. <https://doi.org/10.1016/j.cobeha.2021.03.008>

- Boyer, T. W., & Levine, S. C. (2015). Prompting children to reason proportionally: Processing discrete units as continuous amounts. *Developmental Psychology*, *51*(5), 615–620. <https://doi.org/10.1037/a0039010>
- Chin, A., Markey, A., Bhargava, S., Kassam, K. S., & Loewenstein, G. (2017). Bored in the USA: Experience sampling and boredom in everyday life. *Emotion*, *17*(2), 359–368. <https://doi.org/10.1037/emo0000232>
- Dailey, R., Romo, L., Myer, S., Thomas, C., Aggarwal, S., Nordby, K., Johnson, M., & Dunn, C. (2018). The buddy benefit: Increasing the effectiveness of an employee-targeted weight-loss program. *Journal of Health Communication*, *23*(3), 272–280. <https://doi.org/10.1080/10810730.2018.1436622>
- de Hevia, M. D., & Spelke, E. S. (2010). Number-space mapping in human infants. *Psychological Science*, *21*(5), 653–660. <https://doi.org/10.1177/0956797610366091>
- Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*(1), 105–115. <https://doi.org/10.1037/h0030644>
- Denison, S., & Xu, F. (2010). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science*, *13*(5), 798–803. <https://doi.org/10.1111/j.1467-7687.2009.00943.x>
- Doan, T., Castro, A., Bonawitz, E., & Denison, S. (2020). “Wow, I did it!”: Unexpected success increases preschoolers’ exploratory play on a later task. *Cognitive Development*, *55*, Article 100925. <https://doi.org/10.1016/j.cogdev.2020.100925>
- Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology*, *5*, Article 36. <https://doi.org/10.3389/fpsyg.2014.00036>
- Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, *1*(4), 324–340. [https://doi.org/10.1016/0010-0285\(70\)90019-8](https://doi.org/10.1016/0010-0285(70)90019-8)
- Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don’t know. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(13), 3492–3496. <https://doi.org/10.1073/pnas.1515129113>
- Gürel, Ç., Brummelman, E., Sedikides, C., & Overbeek, G. (2020). Better than my past self: Temporal comparison raises children’s pride without triggering superiority goals. *Journal of Experimental Psychology: General*, *149*(8), 1554–1566. <https://doi.org/10.1037/xge0000733>
- Gweon, H., Asaba, M., & Bennett-Pierre, G. (2017). Reverse-engineering the process: Adults and preschoolers’ ability to infer the difficulty of novel tasks. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th annual conference of the Cognitive Science Society* (pp. 458–463). Cognitive Science Society.
- Hembacher, E., & Ghetti, S. (2014). Don’t look at my answer: Subjective uncertainty underlies preschoolers’ exclusion of their least accurate memories. *Psychological Science*, *25*(9), 1768–1776. <https://doi.org/10.1177/0956797614542273>
- Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial location in young children. *Cognitive Psychology*, *27*(2), 115–147. <https://doi.org/10.1006/cogp.1994.1014>
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, *7*(5), Article e36399. <https://doi.org/10.1371/journal.pone.0036399>
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development*, *85*(5), 1795–1804. <https://doi.org/10.1111/cdev.12263>
- Kim, K., Sano, M., De Freitas, J., Haber, N., & Yamins, D. (2020). Active world model learning with progress curiosity. *ArXiv*, 2007.07853 [Cs, Stat]. <http://arxiv.org/abs/2007.07853>
- Leonard, J. A., Lee, Y., & Schulz, L. E. (2017). Infants make more attempts to achieve a goal when they see adults persist. *Science*, *357*(6357), 1290–1294. <https://doi.org/10.1126/science.aan2317>
- Leonard, J. A., Sandler, J., Nerenberg, A., Rubio, A., Schulz, L. E., & Mackey, A. P. (2020). Preschoolers are sensitive to their performance over time. *Proceedings of the 42st Annual Conference of the Cognitive Science Society*.
- Lucca, K., Horton, R., & Sommerville, J. A. (2020). Infants rationally decide when and how to deploy effort. *Nature Human Behaviour*, *4*(4), 372–379. <https://doi.org/10.1038/s41562-019-0814-0>
- Martin, A., Ryan, R. M., & Brooks-Gunn, J. (2013). Longitudinal associations among interest, persistence, supportive parenting, and achievement in early childhood. *Early Childhood Research Quarterly*, *28*(4), 658–667. <https://doi.org/10.1016/j.ecresq.2013.05.003>
- Messer, D. J., McCarthy, M. E., McQuiston, S., MacTurk, R. H., Yarrow, L. J., & Vietze, P. M. (1986). Relation between mastery behavior in infancy and competence in early childhood. *Developmental Psychology*, *22*(3), 366–372. <https://doi.org/10.1037/0012-1649.22.3.366>
- Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*(1), 19–46. <https://doi.org/10.1007/s11409-013-9094-7>
- Metcalfe, J., & Jacobs, W. J. (2010). People’s study time allocation and its relation to animal foraging. *Behavioural Processes*, *83*(2), 213–221. <https://doi.org/10.1016/j.beproc.2009.12.011>
- Metcalfe, J., & Kornell, N. (2005). A Region of Proximal Learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463–477. <https://doi.org/10.1016/j.jml.2004.12.001>
- Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. *Science*, *244*(4907), 933–938. <https://doi.org/10.1126/science.2658056>
- Newcombe, N. S. (2014). The origins and development of magnitude estimation. *Ecological Psychology*, *26*(1–2), 147–157. <https://doi.org/10.1080/10407413.2014.875333>
- Niebaum, J., & Munakata, Y. (2020). Deciding what to do: Developments in children’s spontaneous monitoring of cognitive demands. *Child Development Perspectives*, *14*(4), 202–207. <https://doi.org/10.1111/cdep.12383>
- O’Leary, A. P., & Sloutsky, V. M. (2019). Components of metacognition can function independently across development. *Developmental Psychology*, *55*(2), 315–328. <https://doi.org/10.1037/dev0000645>
- Oudeyer, P., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, *11*(2), 265–286. <https://doi.org/10.1109/TEVC.2006.890271>
- Parsons, J. E., & Ruble, D. N. (1977). The development of achievement-related expectancies. *Child Development*, *48*(3), 1075–1079. <https://doi.org/10.2307/1128364>
- Payne, S. J., Duggan, G. B., & Neth, H. (2007). Discretionary task interleaving: Heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology: General*, *136*(3), 370–388. <https://doi.org/10.1037/0096-3445.136.3.370>
- Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, *102*(3), 531–549. <https://doi.org/10.1037/a0019243>
- Renfree, I., Harrison, D., Marshall, P., Stawarz, K., & Cox, A. (2016). Don’t kick the habit: The role of dependency in habit formation apps. *CHI EA ’16 proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems* (pp. 2932–2939). Association for Computing Machinery (ACM). <https://doi.org/10.1145/2851581.2892495>
- Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science*, *1*(2), 291–297. <https://doi.org/10.1111/1467-7687.00044>

- Schulz, L. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences*, *16*(7), 382–389. <https://doi.org/10.1016/j.tics.2012.06.004h>
- Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development*, *84*(2), 443–454. <https://doi.org/10.1111/j.1467-8624.2012.01865.x>
- Siegel, M. H., Magid, R. W., Pelz, M., Tenenbaum, J. B., & Schulz, L. E. (2021). Children’s exploratory play tracks the discriminability of hypotheses. *Nature Communications*, *12*(1), Article 3598. <https://doi.org/10.1038/s41467-021-23431-2>
- Siemens, J. C., Smith, S., Fisher, D., Thyroff, A., & Killian, G. (2015). Level up! The role of progress feedback type for encouraging intrinsic motivation and positive brand attitudes in public versus private gaming contexts. *Journal of Interactive Marketing*, *32*, 1–12. <https://doi.org/10.1016/j.intmar.2015.07.001>
- Son, L. K., & Sethi, R. (2006). Metacognitive control and optimal learning. *Cognitive Science*, *30*(4), 759–774. [https://doi.org/10.1207/s15516709cog0000\\_74](https://doi.org/10.1207/s15516709cog0000_74)
- Son, L. K., & Sethi, R. (2010). Adaptive learning and the allocation of time. *Adaptive Behavior*, *18*(2), 132–140. <https://doi.org/10.1177/1059712309344776>
- Stipek, D. J., & Hoffman, J. M. (1980). Development of children’s performance-related judgments. *Child Development*, *51*(3), 912–914. <https://doi.org/10.2307/1129485>
- Stipek, D. J., Roberts, T. A., & Sanborn, M. E. (1984). Preschool-age children’s performance expectations for themselves and another child as a function of the incentive value of success and the salience of past performance. *Child Development*, *55*(6), 1983–1989. <https://doi.org/10.2307/1129773>
- Ten, A., Kaushik, P., Oudeyer, P.-Y., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*, *12*(1), 5972. <https://doi.org/10.1038/s41467-021-26196-w>
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285. <https://doi.org/10.1126/science.1192788>
- Wang, R., Lehman, J., Rawal, A., Zhi, J., Li, Y., Clune, J., & Stanley, K. O. (2020). Enhanced POET: Open-ended reinforcement learning through unbounded invention of learning challenges and their solutions. *ArXiv:2003.08536* [Cs]. <http://arxiv.org/abs/2003.08536>
- Xia, M., Poorthuis, A. M. G., Zhou, Q., & Thomaes, S. (2021). Young children’s overestimation of performance: A cross-cultural comparison. *Child Development*, *93*(2), e207–e221. <https://doi.org/10.1111/cdev.13709>

Received March 1, 2022

Revision received July 5, 2022

Accepted August 12, 2022 ■