Title

Developmental changes in children's training strategies

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Serko, Daniil
Leonard, Julia Anne
Ruggeri, Azzurra

Publication Date

2022

Developmental changes in children's training strategies

Daniil Serko (serko@mpib-berlin.mpg.de)
MPRG iSearch, Max Planck Institute for Human Development

Julia Leonard (julia.leonard@yale.edu)
Department of Psychology, Yale University

Azzurra Ruggeri (ruggeri@mpib-berlin.mpg.de)
MPRG iSearch, Max Planck Institute for Human Development
Department of Cognitive Science, Central European University
School of Social Sciences and Technology, Technical University Munich

Abstract

Effective practice is key to learning. Yet, it is unclear whether young children have the ability to make effective and adaptive training choices. In this project, we investigated 4- to 7-year-old children's (n=146) ability to tailor their training strategies to optimize performance outcomes. Children were presented with one easy and one difficult guessing game and were asked to choose which game they wanted to practice. Crucially, before they chose, they were told that they would eventually be tested either on the game of their choice (*Choice* condition) or on the game the computer would randomly pick (*Random* condition). Contrary to our hypotheses, we found that condition *per se* did not predict children's training choices. However, we found that older children were more likely to make *effective* and *adaptive* training choices than younger children. Overall, our results indicate that children's training choices improve from ages 4 to 7 and inform the development of interventions to support strategic learning.

Keywords: active learning, calibration, study-effort allocation, metacognition, development, decision making

Introduction

Imagine you have a Math test and a French test tomorrow, but you only have one hour to study. Should you split your time equally between the two subjects or devote the full hour to focus on only one of them? If you are better at French than Math and want to do well on both tests, it would probably make sense to devote most of your time to studying Math. However, if you only care about your French test, you might want to devote the hour to perfecting your French. As this example illustrates, effectively deciding how to allocate your limited resources depends on your *abilities* and *goals*. Although deciding whether to study Math or French may seem trivial (it is probably okay to fail one Math or French test), decisions about where and how we spend our efforts more generally build up over time to determine what we learn and who we become. Thus, understanding how young children make such decisions early in life may be especially critical to informing interventions supporting children's later learning and development.

Past work has shown that adults are adept at allocating resources based on their abilities and goals to optimize performance. For example, Ten, Kaushik, Oudeyer, and Gottlieb (2021) presented participants with games that varied in difficulty and told them to play for a given number of trials. With no external constraints, adults spent their time playing easier games, on which they made fast progress (Baranes, Oudeyer, & Gottlieb, 2014). However, when participants were instructed to learn all games because they would eventually be tested on them, they were more likely to spend their time playing more difficult games (Ten et al., 2021). Similarly, when preparing for a test of novel word pairs, adults studied items of intermediate difficulty and avoided spending time studying items they already knew or that were very difficult (Metcalfe & Kornell, 2005). Furthermore, individual differences in adults' ability to monitor and track their performance over time—that is, their *metacognition*—relate to their test outcomes (Dunlosky & Nelson, 1992; Metcalfe & Kornell, 2005), indicating that having a fine-tuned awareness of one's own abilities and learning trajectory supports performance outcomes. However, it is unclear when this ability to effectively and adaptively allocate one's effort and resources (e.g., time) emerges and how it develops across childhood. In this project, we ask whether 4- to 7-year-olds can make effective decisions about what to practice based on their own abilities and given goals in order to maximize expected rewards.

Prior work suggests that even toddlers have some awareness of their abilities. For example, 20-month-olds selectively ask caregivers for help when they are unsure of the location of a hidden toy (Goupil, Romand-Monnier, & Kouider, 2016) and preschoolers are more confident about memory items that they answer correctly versus incorrectly (Hembacher & Ghetti, 2014). Young children can also use cues related to physical features of the task, social learning, and prior performance to infer their own and others' performance. For example, Bridgers, Jara-Ettinger, and Gweon (2020) showed that children infer task difficulty from physical features of the task, like the number of buttons on a toy, and use this information to decide which game to teach another person (see also Magid, DePascale, & Schulz, 2018). Moreover, infants and preschoolers put more effort into achieving a goal when they see an adult working hard versus effortlessly succeeding at the same or different goal (Leonard, Lee, & Schulz, 2017; Leonard et al., 2020; Lucca, Horton, & Sommerville, 2020). There is also evidence that children adapt their efforts based on external goals. For example, preschoolers engage in fewer goal-directed and more playful actions when they are told to

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

play versus achieve a specific goal (Chu & Schulz, 2020), and are less likely to explore a novel toy after one specific affordance is explicitly demonstrated (Bonawitz et al., 2011).

However, prior work also suggests that these inferences, and children's ability to *act* on them, go through substantial change across development (Metcalfe & Finn, 2013). First, a large body of work has shown that children often tend to overestimate their physical and mental abilities (Flavell, Friedrichs, & Hoyt, 1970; Plumert, 1995; Schneider, 1998; Stipek & Hoffman, 1980), even when provided with explicit feedback on their performance (O'Leary & Sloutsky, 2019). Second, although children may have some general awareness of their abilities, they do not always act based on this information to optimize their resource allocation. For example, 4- to 5-year-old children do not allocate more study time to a memory task they previously did poorly on (Flavell et al., 1970). Similarly, third graders do not spend extra study time on items that they previously answered incorrectly on a memory test, even though they report being less confident about their knowledge of these items(Metcalfe & Finn, 2013). It is not until fifth grade that students behave like adults and devote more study time to the yet-to-be-learned items (Metcalfe & Finn, 2013). Moreover, when told to choose which game to play with in the absence of a test, 11-year-old children and adults, but not 6-year-olds, prefer to play with a less cognitively demanding game (Niebaum, Chevalier, Guild, & Munakata, 2019).

To summarize, unlike adults and older children, younger children do not invest resources towards compensating for their weaknesses before a test and do not devote time to easier tasks in the absence of a test. In this sense, younger children do not seem to be making effective (choosing a strategy to optimize performance given external goals) or adaptive (flexibly changing strategies across goal contexts) decisions about effort allocation. This pattern of results may be driven by developmental changes in children's beliefs about their abilities (Schneider, 1998), executive function and working memory (for a recent review, see Roebers (2017); Marulis, Baker, and Whitebread (2020)), their ability to accurately monitor task demands (Chevalier, Martis, Curran, & Munakata, 2015), or adapt to specific task features (Lindow & Betsch, 2018; Meder, Wu, Schulz, & Ruggeri, 2019). However, past work has not compared children's learning choices across different explicit goals *within or between* participants. Thus, we do not have definitive evidence of whether young children can adapt their training strategies to make effective choices when facing distinct goals, even on simple tasks.

The present study

In this study, we investigate whether 4- to 7-year-old children make effective and adaptive training choices to optimize their performance given specific task goals. To this end, we created two novel, interactive, and child-friendly games where children had to guess 8 different cards (animals or objects) from their pictures (game 1) or sounds (game 2). Children were familiarized with 4 cards from each game to demonstrate that one of the games was very easy (all cards could be correctly guessed) and one was very difficult (none of the cards could be guessed correctly; stimuli were selected after extensive pilot testing). Children were told whether their guesses were correct or incorrect, but they did not receive corrective feedback for unknown cards. Children were randomly assigned to one of two conditions: In the *Choice* condition, we told them that they could eventually choose which of the two games to be tested on, whereas in the *Random* condition, we told them that the computer would randomly chose one of the two games for them to be tested on. Children were told that they would get stickers based on their performance at test, which would include all the 8 cards of a game. They were then asked to choose which game they wanted to practice, that is, which set of 4 non-familiarization cards they wanted to learn about before the test.

The optimal strategy in the *Choice* condition is to practice the easy game and then choose it at test. This approach guarantees to score 8 out of 8 points—children already succeeded on the 4 familiarization cards of the easy game and can then learn about the 4 non-familiarization cards. If they chose to get tested on the difficult game, they would not be able to score more than 4 points in any case, as they would not get the chance to learn the correct answer for the 4 familiarization cards of the difficult game they had initially failed to guess. In the *Random* condition, the optimal strategy is to train the difficult game, as it ensures achieving at least 4/8 points, whatever game is presented at test. If children decided to train the easy game in the *Random* condition and happened to be tested on the difficult game, they would end up scoring zero points. We, therefore, predicted that children would behave optimally and make effective training choices, choosing to train and test on the easy game in the *Choice* condition and training the difficult game in the *Random* condition. However, in light of the developmental literature reviewed above, we also predicted age-related improvements in children's ability to make effective training choices, adapting their effort-allocation decisions to the given goal.

Children played two rounds of the game, with condition assignment manipulated both within and between participants. The resulting 2x2 design allowed us to compare group differences and further investigate, at the individual level, children's adaptiveness in two ways: 1) children's ability to make effective training choices *across rounds*, that is flexibly changing (or not) their training strategy from round 1 to round 2 depending on the given goals, and 2) children's ability to learn from their mistakes, that is, to adjust their training strategies after failure. We predicted that children's ability to make adaptive training decisions across goals and to learn from their mistakes would also increase with age.

Methods

Participants Participants were 146 4- to 7-year-old children (87 female, $M = 70.30$ months; $SD = 12.92$ months; Range: 48 to 95 months). We recruited and tested participants in the public Zoo in Berlin, Germany. Fifteen additional children were tested but dropped from further analysis due to failure to answer the comprehension questions correctly (N = 6; see below), parental interference (N = 2), tablet malfunction (N = 3), or because they were outside our age range (N = 4).

The sample size was calculated based on a power analysis performed on a simulated dataset. According to this analysis, we would have needed to collect data on at least 132 children in total to be able to detect a difference between conditions with power = .90. We tested a few additional children to ensure a more even age distribution in our sample. Before starting the experimental session, parents signed written informed consent, and children were asked to give verbal consent to participate. The study was approved by the ethics committee of the Max Planck Institute for Human Development, Berlin.

Materials & Design Participants were introduced to two guessing games on a touchscreen tablet. The tablet presented a split-screen with a *Pictures* game (8 cards) on one side and a *Sounds* game (8 cards) on the other side (see Fig. 1). Children were told that in the Pictures game, they had to guess the animal (or object) illustrated on a picture, whereas in the Sounds game, they had to guess which object (or animal) had generated the presented sound. We ensured through pilot testing (N = 16, $M = 74.19$ months) that one of the two games would be extremely easy (i.e., presenting sounds of familiar animals or objects, such as a cat or a bell, that all children could correctly guess), and the other extremely difficult (e.g., presenting pictures of unfamiliar animals or objects, such as a pangolin or a kitchen-mandolin, that no children in our pilot sample could correctly guess). Note that, in our final sample, children failed 1/4 familiarization cards of the easy game in 32 rounds (11%) and children correctly guessed 1/4 familiarization cards of the difficult game in 8 rounds (3%). We decided to include all data, including those from these rounds, in the analyses reported below. Note that none of the results drastically change if we exclude these rounds from the analyses. We counterbalanced whether the Pictures or Sounds games were easy or difficult and whether they were assigned the animals or objects stimuli. The experimental session included two rounds, each consisting of three phases: familiarization, training, and test.

In the familiarization phase, children could click on the top 4 cards of the Pictures game and the top 4 cards of the Sounds game. Children were told whether they guessed correctly or not but were not given the correct answers for the cards they guessed wrong or did not know (see Fig. 1). Children then completed a comprehension check where they indicated which game they performed better on. Only children who answered the comprehension check correctly were included in the analyses (see above). After the familiarization phase, children were told that they would eventually be tested on all of the 8 cards of one of the two games (half of which they had already seen). Children were told that they would win a sticker for each correct answer and lose one sticker for each incorrect answer at test, which would include all the 8 cards of a game. Crucially, we either told children that they could decide which game they would eventually be tested on (*Choice* condition) or that the computer had already randomly selected which game they would be tested on but that we would not know which game that was is until the test phase (*Random* condition).

In the training phase, children had to choose which game they wanted to practice, that is, which set of 4 non-familiarization cards they wanted to learn about before the test. We then asked children to guess the 4 non-familiarization cards, and, unlike in the familiarization phase, we provided them with the correct answers for the animals or objects they could not guess correctly. Children then proceeded to the test phase, where they guessed all eight cards of one game. To maximize hypothesized differences across conditions, children in the *Random* condition were always tested on the difficult game.

Children played a second round of the game that was identical to the first round but with new stimuli. Children were again randomly assigned to either the *Choice* or the *Random* condition, overall resulting in a 2x2 within/between-subjects design. After both rounds were completed, children received the stickers they won and were thanked for participation.

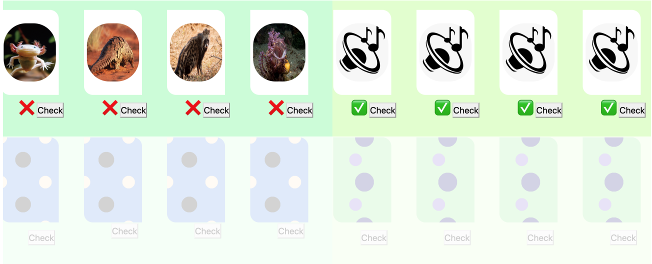

Figure 1: Example of the calibration phase.

Results

Training choice We first examined whether participants tailored their training choices to the given goals in round 1, when they had no prior experience with the task. In the *Choice* condition, 39 out of 72 children chose to train the easy game (54%; $p = .278$, binomial test), and in the *Random* condition, 33 out of 74 children chose to train the difficult game (45%; $p = .852$, binomial test). A logistic regression predicting training choices with condition (*Choice* vs. *Random*) revealed that condition did not predict children's training choices in round 1 ($p = .880$, $OR = 1.051$ [0.548 – 2.018]). Adding age to the model revealed a significant, positive effect of age on children's training choices ($p = .048$, $OR = 0.964$ [0.930 – 1.000]). Specifically, older children preferred to train the difficult game across conditions, while younger children preferred to train the easy game across conditions. There was no age by condition interaction on training choices.

We found similar results in round 2. In the *Choice* condition, 50 out of 76 children chose to train the easy game

(66%; $p = .004$, binomial test), and in the *Random* condition, 29 out of 70 children chose to train the difficult game (41%; $p = .940$, binomial test). Again, condition did not significantly predict children's training choices ($p = .369$, $OR = 0.735$ [0.374 – 1.438]). As in round 1, adding age to the model revealed a significant, positive effect of age on children's training choices ($p = .006$, $OR = 0.942$ [0.924 – 1.044]). There was no age by condition interaction on training choices.

Examining both rounds with age, condition, and round as predictors, we found that only age ($p = .048$, $OR = 0.964$ [0.928 – 0.999]), but not condition ($p = .399$, $OR = 5.505$ [0.106 – 337.770]) nor round ($p = .294$, $OR = 8.799$ [0.160 – 569.010]), predicted training choices. We found no interaction effects of predictors on training choice.

Test performance We ran a linear regression to examine how condition (*Choice* vs. *Random*) predicted children's performance at test (that is, how many items they guessed correctly) across both rounds. We found that condition was a significant predictor (B = -0.863, $p < .001$) of participants performance. As expected, children obtained more points in the *Choice* condition ($M = 6.96$, $SD = 2.09$) compared to the *Random* condition ($M = 1.12$, $SD = 1.21$; $t(236.53) = 29.30$, $p < .001$, Welch two-sample t-test). Adding age to the model revealed no further significant effects on children's test performance. The effect of condition on test performance (B = -1.064, $p < .001$) also held when adding age and round (round 1 vs. round 2) as a predictor.

Effectiveness of training choice The fact that condition (or its interaction with age) did not predict children's training choices goes against our hypotheses. However, this apparently puzzling behavior makes more sense when we consider that in a considerable proportion of *Choice*-condition rounds (27 out of 148 rounds; 18%), those children who chose to train the difficult game also decided to get tested on the easy game and obtained the maximum score, see Fig. 3). The learning choices of these children (1 4-year-old; 5 5-year-olds, 13 6-year-olds, and 8 7-year-olds) are in fact, sophisticated, and cannot be considered ineffective—as they still led to maximum scores and did not imply any loss of resources. Therefore, to analyze the *effectiveness* of participants' training choices, instead and beyond their training choices, we recoded them as *effective* in the following cases:

- train the easy game in the *Choice* condition (Round 1: n = 39, Round 2: n = 50, total n = 89)
- train the difficult game in the *Random* condition (Round 1: n = 33, Round 2: n = 29, total n = 62)
- train the difficult game in the *Choice* condition, then choose the easy game at test and achieve the maximum score (8 points; Round 1: n = 13, Round 2: n = 14, total n = 27)

Participants' choices were coded as *ineffective* otherwise (Round 1: n = 61, Round 2: n = 53, total n = 114). Overall, children who made effective training choices achieved more points ($M = 5.51$, $SD = 3.00$) than children who made ineffective training choices ($M = 1.85$, $SD = 2.70$; $t(259.08) = 10.806$, $p < .001$).

We first examined whether participants made effective training choices when they had no prior experience with the task, in round 1. A logistic regression revealed a significant effect of condition on training effectiveness ($p < .001$, $OR = 0.310$ [0.155 – 0.618]), with children making more effective training choices in the *Choice* condition than in the *Random* condition. When we added age to the model, condition remained significant ($p = .010$, $OR = 0.004$ [0.000 – 0.267]), but age did not predict children's effective training choices ($p = .995$, $OR = 1.000$ [0.962 – 1.039]). The interaction of age and condition significantly predicted participants training choices ($p = .040$, $OR = 1.061$ [1.002 – 1.124]): As can be seen in Figure 3, younger children were as effective as older children in the *Choice* condition. However, in the *Random* condition, we see clear improvement in children's ability to make effective training choices with age.

We found similar results for round 2. A logistic regression predicting effective training by condition revealed a significant effect of condition on training effectiveness ($p < .001$, $OR = 0.133$ [0.061 – 0.289]). When we added age to the model, condition remained significant ($p < .001$, $OR = 0.000$ [0.000 – 0.011]), while age did not predict children's effective training choices ($p = .322$, $OR = 0.975$ [0.927 – 1.025]). The interaction of age and condition again significantly predicted participants training choices ($p = .002$, $OR = 1.108$ [1.037 – 1.184]).

A logistic regression predicting effective training across both rounds with age, condition, and round as predictors revealed no significant effect of age ($p = .396$, $OR = 0.998$ [0.993 - 1.003]). However, condition ($p < .001$, $OR = 0.002$ [0.002 – 0.002]), round ($p < .001$, $OR = 9.562$ [9.489 – 9.634]), the interaction of age and condition ($p < .001$, $OR = 1.070$ [1.063 – 1.079]), the interaction of age and round ($p < .001$, $OR = 0.979$ [0.974 – 0.985]), the interaction of condition and round ($p < .001$, $OR = 0.030$ [0.030 – 0.030]), and the interaction of the three predictors ($p < .001$, $OR = 1.0362$

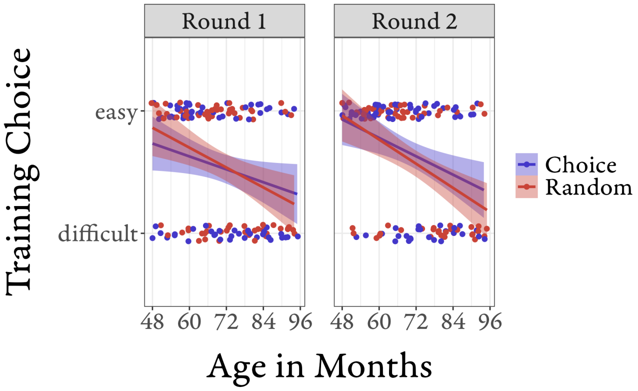

Figure 2: Participants training choices by condition and age in months in round 1 and round 2.

[1.030 – 1.043]) all significantly predicted participants' training choice effectiveness. These results echo those from the analyses reported above for round 1 and round 2: Although younger children were as effective as older children in the *Choice* condition, in the *Random* condition, we see a clear age improvement in children's ability to make effective training choices. The three-way interaction of condition, age, and round further indicates that the interaction between age and condition is more pronounced in round 2 compared to round 1 (see Fig. 3).

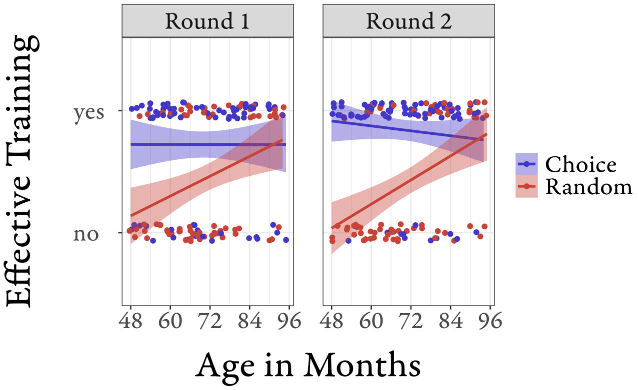

Figure 3: Participants effective training choices by condition and age in months in rounds 1 and 2.

Adaptiveness across rounds Here, we examined children's adaptiveness - their ability to make effective training choices *across rounds*, flexibly changing (or not) their training strategy from round 1 to round 2 depending on the given goals (see Fig. 4). When considering only the first round, 85 out of 146 children (58%; $p = .028$, binomial test) made an effective training choice. Most of the children who made an effective choice in round 1 also made an effective choice in round 2 (59 out of 85 children, 69%; $p < .001$, binomial test with chance = 25%). That is, overall, 59 out of 146 participants (40%) made an effective training choice in both rounds ($p < .001$, binomial test with chance = 25%). A logistic regression revealed that age positively predicted children's ability to make effective training choices across the two rounds ($p < .001$, $OR = 1.037$ [1.018 – 1.057]). We did not find an effect of condition combination (same condition across rounds or different condition across rounds) on children's adaptiveness (model controlled for age).

Discussion

In this study, we investigated 4- to 7-year-old children's ability to tailor their training strategies to a given goal in order to maximize expected rewards. We found that, with age, children were better able to make effective and adaptive training choices across different goals (*Choice* vs. *Random* condition). Specifically, younger children preferred to practice the easy game, and older children generally preferred to practice the difficult game irrespective of the goal.

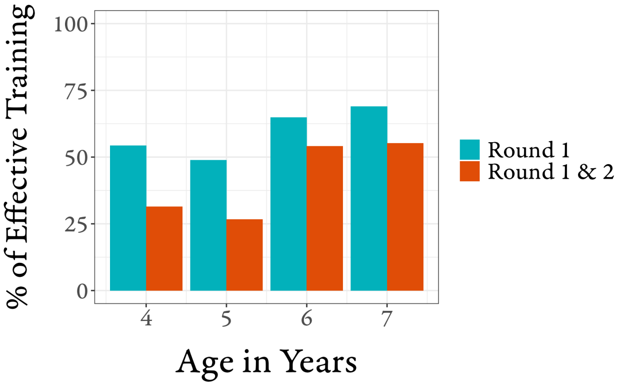

Figure 4: Percentage of children who made an effective training choice at least in round 1 and an effective training choice in rounds 1 and 2.

There are several possible explanations for young children's preference for the easy game. One possibility is that younger children just did not appreciate the difference in difficulty between tasks. We know from previous work that being aware, monitoring, and updating information about one's previous performance is cognitively demanding (Legare, Mills, Souza, Plummer, & Yasskin, 2013). Previous research has highlighted a connection between children's efficient goal maintenance and their working memory capacities (Marcovitch, Boseovski, Knapp, & Kane, 2010). These metacognitive skills improve over the course of development (Diamond, 2013), supporting and potentially constraining children's performance in effort-allocation tasks, such as the one presented here. However, we are fairly confident that children in our study were aware of the difference in difficulty between the presented games. First, all children included in this study passed the comprehension questions explicitly addressing this point. Second, the previous work reviewed in the introduction suggests that even toddlers have some awareness of their own abilities and can, to a certain extent, track them across tasks and over time. Finally, we made sure that children could easily track their past performance by providing explicit memory aids throughout the game.

An alternative interpretation is that young children did appreciate relative task difficulty yet failed to plan their efforts accordingly (Bennett-Pierre, Asaba, & Gweon, 2018; Gweon, Asaba, & Bennett-Pierre, 2017; Metcalfe & Finn, 2013). On the one hand, young children in the *Random* condition might have engaged in wishful thinking, hoping (or even being somehow confident) that they will eventually get tested on the easy game. This mindset would be in line with previous developmental work, showing that children are often overly positive about future outcomes when they expect them to be beneficial for themselves (Bernard, Clement, & Mercier, 2016). On the other hand, it is possible that young children found that playing the difficult game was frustrating and just *not fun*, and therefore avoided playing it, even if that meant potentially failing at test. This finding is in line with previous results documenting that fifth and sixth graders generally prefer tasks labeled as "easy" over those labeled as "difficult" (Hom & Maxwell, 1983). It could also just be that young children did not clearly understand the game instructions. For example, they may not have realized that the fact that the "computer would choose which game they would be tested on" in the *Random* condition meant that there was a 50% probability of eventually being tested on the difficult game.

Future work should control for these potential confounds by i) making the instructions more explicit, ii) introducing check questions—for example, asking them how likely the computer is to eventually test them in the two games—, iii) introducing stronger performance incentives that would motivate children to practice the difficult game even if frustrating, or iv) by contrasting, instead of a *Random* and a *Choice* condition, an *Easy*, a *Difficult*, and a *Random* condition, in which children are clearly explained (and illustrated) before training that they will be tested on the easy game, the difficult game, or on one of the two games at random. Finally, v) future work might want to connect these research questions and findings more explicitly to the developmental trajectory of executive function and metacognitive skills across childhood.

Older children's overall preference to practice the difficult game across conditions is in line with several previous studies reporting participants' reluctance to invest resources in learning already-familiar items (Masur, McIntyre, & Flavell, 1973; Metcalfe & Kornell, 2005). Indeed, we found that older children practicing the difficult game in the *Choice* condition did not necessarily demonstrate a sub-optimal training approach. In a considerable proportion of *Choice*-condition rounds (19%), after the familiarization phase, older children were rightly confident in their abilities to guess the easy-game cards without training. Consequently, their decision to practice the difficult game—which may have been more appealing, as it offered the opportunity to learn something new and therefore was potentially more fun—did not prevent them from obtaining the maximum score when they chose to test on the easy game, although they had not practiced its non-familiarization cards. In this sense, their training choices were, in fact, pretty sophisticated and equally effective.

Considering this unexpected though sensible behavior, one way to obtain the pattern of results we had initially hypothesized could be to modify the familiarization phase to induce a stronger difference in expected pay-offs between the two training strategies across conditions. For example, the familiarization of the easy game could include three easy cards and one difficult card, thus reducing children's confidence that they would be able to obtain the maximum score in the easy game without practice. Another possibility would be to make errors in the easy game more costly, for example, by establishing the rule that one needs to guess all cards of one game correctly to win stickers.

Our results also demonstrate children's improving ability to adapt their effort-allocation and training strategies to given goals. Indeed, older children are not only more likely than younger children to make effective training choices by tailoring their strategies to the given goal but also more likely to dynamically adapt their strategies across rounds, flexibly changing (or not) their training strategy from round 1 to round 2 if necessary. This is in line with recent work indicating a developmental increase from age 4 to 10 years in children's ability to adapt their pre-decisional and decision-making strategies to the characteristics of a given task (for a review, see De Simone and Ruggeri (2022)).

To conclude, our study indicates that from age 4 to 7, children substantially improve in their ability to make effective, adaptive effort-allocation, training choices, which appear to be often pretty sophisticated already by age 7. These results might have implications for the development of interventions that scaffold children's effective study strategies, in this sense supporting their later independent learning more generally. In particular, our results suggest that young children may need more explicit guidance about how to best allocate their training efforts to compensate for their weaknesses versus boosting their strengths.

References

Baranes, A. F., Oudeyer, P.-Y., & Gottlieb, J. (2014). The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Frontiers in Neuroscience*, *8*, 317.

Bennett-Pierre, G., Asaba, M., & Gweon, H. (2018). Preschoolers consider expected task difficulty to decide what to do and whom to help. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*.

Bernard, S., Clement, F., & Mercier, H. (2016, January). Wishful thinking in preschoolers. *Journal of Experimental Child Psychology*, *141*, 267–274.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011, September). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–330.

Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2020). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour*, *4*(2), 144–152.

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience*, *27*(6), 1125–1136.

Chu, J., & Schulz, L. (2020). Exploratory play, rational action, and efficient search. *Proceedings of the 42st Annual Conference of the Cognitive Science Society*.

De Simone, C., & Ruggeri, A. (2022). Searching for information, from infancy to adolescence. In I. C. Dezza, E. Schulz, & C. M. Wu (Eds.), *The Drive for Knowledge: The Science of Human Information-Seeking*. (pp. 77–100). Cambridge, MA, US: Cambridge University Press.

Diamond, A. (2013). Executive functions. *Annual Review of*

Psychology, *64*, 135–168.

Dunlosky, J., & Nelson, T. O. (1992, July). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, *20*(4), 374–380.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970, October). Developmental changes in memorization processes. *Cognitive Psychology*, *1*(4), 324–340.

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*, *113*(13), 3492–3496.

Gweon, H., Asaba, M., & Bennett-Pierre, G. (2017). Reverse-engineering the process: Adults' and preschoolers' ability to infer the difficulty of novel tasks. *Proceedings of the 37th Annual Conference of the 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.

Hom, H. L., & Maxwell, F. R. (1983, March). The impact of task difficulty expectations on intrinsic motivation. *Motivation and Emotion*, *7*(1), 19–24.

Legare, C. H., Mills, C. M., Souza, A. L., Plummer, L. E., & Yasskin, R. (2013, January). The use of questions as problem-solving strategies during early childhood. *Journal of Experimental Child Psychology*, *114*(1), 63–76.

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.

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*.

Lindow, S., & Betsch, T. (2018). Child decision-making: On the burden of predecisional information search. *Journal of Cognition and Development*, *19*(2), 137–164.

Lucca, K., Horton, R., & Sommerville, J. A. (2020, April). Infants rationally decide when and how to deploy effort. *Nature Human Behaviour*, *4*(4), 372–379.

Magid, R. W., DePascale, M., & Schulz, L. E. (2018). Four- and 5-year-olds infer differences in relative ability and appropriately allocate roles to achieve cooperative, competitive, and prosocial goals. *Open Mind*, *2*(2), 72–85.

Marcovitch, S., Boseovski, J. J., Knapp, R. J., & Kane, M. J. (2010). Goal neglect and working memory capacity in 4-to 6-year-old children. *Child development*, *81*(6), 1687–1695.

Marulis, L. M., Baker, S. T., & Whitebread, D. (2020). Integrating metacognition and executive function to enhance young children's perception of and agency in their learning. *Early Childhood Research Quarterly*, *50*, 46–54.

Masur, E. F., McIntyre, C. W., & Flavell, J. H. (1973, April). Developmental changes in apportionment of study time among items in a multitrial free recall task. *Journal of Experimental Child Psychology*, *15*(2), 237–246.

Meder, B., Wu, C. M., Schulz, E., & Ruggeri, A. (2019). Development of directed and random exploration in children. *Developmental Science*, e13095.

Metcalfe, J., & Finn, B. (2013, April). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*(1), 19–46.

Metcalfe, J., & Kornell, N. (2005). A Region of Proximal Learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463–477.

Niebaum, J. C., Chevalier, N., Guild, R. M., & Munakata, Y. (2019, February). Adaptive control and the avoidance of cognitive control demands across development. *Neuropsychologia*, *123*, 152–158.

O'Leary, A. P., & Sloutsky, V. M. (2019). Components of metacognition can function independently across development. *Developmental Psychology*, *55*(2), 315.

Plumert, J. M. (1995). Relations between children's overestimation of their physical abilities and accident proneness. *Developmental Psychology*, *31*(5), 866–876.

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental review*, *45*, 31–51.

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science*, *1*(2), 291–297.

Stipek, D. J., & Hoffman, J. M. (1980). Development of children's performance-related judgments. *Child Development*, 912–914.

Ten, A., Kaushik, P., Oudeyer, P.-Y., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*, *12*(1), 1–10.