Child Development

CHILD DEVELOPMENT

EMPIRICAL ARTICLE OPEN ACCESS

Multiple Practice Success Scaffolds Long-Term Test-Enhanced Learning in Preschoolers

Tamás Káldi[1,2] | Ágnes Szőllősi[3,4,5] | Mihály Racsmány[3,4,5]

[1]HUN-REN Hungarian Research Centre for Linguistics, Psycho- and Neurolinguistics Research Group, Budapest, Hungary | [2]Department of Speech and Language Therapy, ELTE Bárczi Gusztáv Faculty for Special Needs Education, Budapest, Hungary | [3]HUN-REN Research Centre for Natural Sciences, Institute of Cognitive Neuroscience and Psychology, Budapest, Hungary | [4]Institute of Psychology, University of Szeged, Szeged, Hungary | [5]Cognitive Medicine Research Group, Competence Centre for Neurocybernetics of the Life Sciences Cluster of the Centre of Excellence for Interdisciplinary Research, Development and Innovation of the University of Szeged, University of Szeged, Szeged, Hungary

Correspondence: Tamás Káldi (kaldi.tamas@nytud.hun-ren.hu; kalditamas@gmail.com)

Received: 22 July 2024 | **Revised:** 28 March 2025 | **Accepted:** 1 July 2025

Funding: Tamás Káldi was supported by the NKFIH project grant No. 143301. Mihály Racsmány was supported by the NAP2022-I-2/2022 Hungarian Brain Research Program 3.0. Ágnes Szőllősi was supported by the Bolyai János Research Scholarship of the Hungarian Academy of Sciences.

Keywords: initial learning | retrieval practice | testing effect

ABSTRACT

Retrieval practice is known to enhance long-term memory retention, a phenomenon termed as retrieval practice effect. Two experiments (NWhite = 202), showed that the effect was present in preschool age (5–6 years) and had a boundary condition, namely, amount of initial learning. Specifically, there was a considerable effect only when children reached a sufficient retrieval success rate during practice as a consequence of multiple initial learning cycles. Corroborating the robustness of our findings, the effect was present for both recall ($d = 0.315$) and recognition ($d = 0.324$) and did not depend on whether the forms of retrieval were the same or different during practice and final test. Important implications for early childhood education and development are discussed.

1 | Introduction

Effective long-term memory formation is crucial for learning and development. A long line of research shows that retrieving to-be-learned material during practice leads to better long-term memory retention than simply repeating it (for overviews, see e.g., Roediger and Karpicke 2006; Rowland 2014). This phenomenon is known as the retrieval practice effect. The effects of retrieving material during practice have been extensively studied in adults; however, our understanding of this mechanism in preschool children remains limited. One question we address is whether it can be observed at all in this age group. This part of our study focuses on replicating earlier results found in the literature (e.g., Fazio and Marsh 2019; Fritz et al. 2007; Haebig et al. 2021; Leonard, Deevy, et al. 2019; Leonard et al. 2020, 2022, 2023; Leonard, Karpicke, et al. 2019) using a large sample of 200 children. The other question we address is a novel and intriguing one and concerns the amount of initial learning. Specifically, we examine how the amount of learning before retrieval practice begins influences the emergence of the effect at hand. This issue is especially important for two reasons. First, there is a dearth of studies on this aspect of the retrieval practice effect in the literature on adults (Karpicke et al. 2014; e.g., Racsmány et al. 2020), and only a few studies on preschoolers. Second, we think that the question has serious pedagogical consequences. Discussions on the retrieval practice effect often emphasize the act of testing itself while neglecting the fact that without a sufficient level of knowledge to test, retrieval success during practice will be low, which can ultimately lead to lower success rates in the long run (for an overview, see Karpicke et al. 2014).

1.1 | Retrieval Practice Effect and the Role of Initial Learning

Retrieval practice refers to repeated retrieval attempts following initial learning. Empirical evidence shows that retrieval practice has a beneficial effect on long-term memory retention compared to the traditional method of learning, whereby material is repeatedly studied multiple times (for an overview see e.g., Karpicke 2017; Roediger and Karpicke 2006). This phenomenon, known as the retrieval practice effect, results in a relatively slow forgetting rate. A few studies showed that in the short term (i.e., when testing happens a few minutes after practice), restudy is more efficient, but retrieval practice leads to more successful retrieval after a few days or weeks (Roediger and Karpicke 2006). However, some investigations have shown the beneficial effect of retrieval practice also in the short term, claiming that null results in these cases might be due to inefficient measurement techniques or low retrieval success during practice (see e.g., Karpicke 2017; Rowland 2014). Taken together, retrieval practice seems to be an efficient method in educational settings, an idea that has been addressed in several studies (Dunlosky et al. 2013; Roediger et al. 2011; Roediger and Karpicke 2006).

The long-term efficiency of retrieval practice is influenced by several factors, such as feedback (Butler and Roediger 2008; Kang et al. 2007), the timing of practice (Pashler et al. 2007; Wheeler and Roediger 1992), or the number of successful retrieval attempts during learning (Kang et al. 2007). The latter factor is a crucial component: if retrieval success does not reach a certain threshold during practice, the effect does not occur (Hogan and Kintsch 1971; Racsmány et al. 2020). However, the role of retrieval success is unfortunately an understudied phenomenon, as recognized by Karpicke et al. (2014): "when studying testing/retrieval practice effects, researchers tend to neglect the importance of initial retrieval success" (p. 243). Additionally, the authors claim that this lack of focus may originate from an emphasis on *testing* itself, and not on instances of successful retrieval, which is also reflected in the terminology. Specifically, Karpicke et al. (2014) suspect that "some researchers may ignore retrieval success problems in retrieval practice experiments because the work is viewed as a 'testing effect' rather than as a retrieval practice effect. The locus of the positive effects on learning, however, is in repeated, successful retrieval" (p. 245). This point is underscored by a limited number of studies that have indeed shown that successful retrieval during practice is essential for the effect to emerge (Karpicke and Roediger 2007b, 2008; Pyc and Rawson 2009).

Some research has reached a different conclusion suggesting that even unsuccessful retrieval attempts during practice may also have a long-term beneficial effect (see e.g., Karpicke 2009; Kornell et al. 2009; Richland et al. 2009). However, it must be emphasized that both the circumstances of these effects and the underlying mechanisms remain topics of debate. One account (Karpicke 2009) proposes that the benefit stems from metacognitive strategies: during unsuccessful attempts, the learner recognizes which items require further study. Hence, on later study trials, they pay special attention to these items, which in turn will be remembered better in the long run. An alternative explanation posits that metacognitive processing may not be necessary. Namely, it is sufficient for learners to recognize the incorrectness of their response and subsequently pay special attention to the correct answer. Supporting this view, research has shown that unsuccessful retrieval attempts followed by feedback enhance learning (Kornell et al. 2009) and that feedback, whether immediate or delayed, mitigates the negative effects of incorrect responses in multiple-choice tests (Butler and Roediger 2008). Similarly, incorrect guessing does not impair long-term retention when feedback is provided (Kang et al. 2011), and feedback only improves retrieval practice efficiency when retrieval is unsuccessful (Pashler et al. 2005). Future research should investigate these accounts and clarify the mechanisms by which unsuccessful retrieval attempts may enhance learning.

Given that such metacognitive skills are likely not yet developed in preschool-aged children, and we lack knowledge regarding how attentional control may work upon detecting a conflict between an incorrectly retrieved item and corrective feedback, we shifted the focus to increasing the rate of successful retrievals during the learning process. This raises a key question: how can we enhance successful retrieval in children of this age? One answer is that a certain level of initial memory performance for the study set has to be reached. In other words, we suggest that with a sufficient level of learning before practice, participants will be more successful at retrieving the material during practice, which, in turn, leads to enhanced memory retention (see e.g., Racsmány et al. 2020). This is one proposal that the present study aims to investigate.

The issue of the amount of initial learning was empirically addressed by, for example, Racsmány et al. (2020). The authors conducted three consecutive experiments that were identical in their procedures: each of them began with an initial learning phase immediately followed by a practice phase, and a final test phase 1 week later. The three experiments differed only in the amount of initial learning: participants were presented with the study material once before the practice phase in the first experiment, three times in the second experiment, and six times in the third experiment. Results showed that this manipulation had an effect: a considerable testing effect was found that was resistant to external effects only in that case when there were multiple initial learning cycles. One interesting finding is that these success rates influenced the presence of the retrieval practice effect. Specifically, in the case of low retrieval success (Experiment 1), the retrieval practice effect reversed due to repeated feedback, resulting in higher success rates on the final test in the restudy condition (see also Pastötter and Bäuml 2016; Storm et al. 2014).

1.2 | Retrieval Practice Effect in Preschoolers

The effect of retrieval practice has been extensively studied in adults; however, relatively few studies have focused on children, particularly preschool children. Therefore, this area of inquiry is interesting for at least three reasons. First, given that retrieval practice has not been as extensively studied in preschoolers as in adults, it is worth investigating the phenomenon and attempting to replicate the effect. Second, there is a dearth of research examining how the amount of learning (i.e., exposure to the study set without retrieval practice) influences the retrieval practice effect in preschool children (see e.g., Gordon and Lowry 2024; Leonard et al. 2020). As the potential role of initial learning has

crucial pedagogical consequences, we consider this question to be an important one. Third, as the mechanisms of memory retrieval in preschool children are different from those in adults, we consider the topic of retrieval practice especially interesting to research in this age group. In general, the process of retrieval is extremely intricate. This complexity is particularly evident in the case of recall, as opposed to recognition, where processes such as searching, evaluation, and monitoring are involved, among others. These processes go through rapid development during the preschool age (for a recent survey, see Newcombe et al. 2022). Hence, studies on the effect of retrieval practice in preschool children are important for investigating whether retrieval is present at all, and if it is, whether it contributes to long-term memory retention.

Regarding the first issue, only a few studies have shown that the effect is present in preschool children. For example, Fritz et al. (2007) conducted two experiments in which the authors investigated the effect of retrieval practice as compared to learning motivated with reward and learning with simple repetition. The age group involved in the experiments ranged from 2.5 to 5 years. In the experiments, children were presented with toy animals and tasked with learning their names (e.g., *Charlie*, *Tinker*). The first experiment included three conditions. In the comparison condition, children were shown the animals along with their names and were required to produce each name upon presentation, followed by a name recall test. The reward condition was identical to the comparison condition, except that children were promised a reward for doing their best. In the expanding retrieval practice condition, the toys and their names were presented, and the names were elicited from the children with feedback on an expanding retrieval schedule. Hence, this condition included two different factors: on the one hand, the factor of the expanding schedule, which may have increased the difficulty of memory retrieval (see e.g., Balota et al. 2011; Karpicke and Roediger 2007a), and feedback, which seems to contribute to the presence of the retrieval practice effect (see e.g., Pashler et al. 2005). The results of the first experiment showed a better memory for the material in the expanding retrieval practice condition, relative to the other two conditions. However, as the authors also admit, the design contained a serious limitation. Namely, in the expanding retrieval practice condition, participants were exposed to the learning material substantially more times than in the other two conditions. Also, the final test was administered shortly after the learning and practice sessions; therefore, the long-term benefit of retrieval practice could not be investigated. In order to overcome these limitations, the authors conducted a second experiment, using two additional conditions in which the amount of exposure to the material was the same as in the expanding retrieval practice condition. This condition was the same as in the first experiment. In the re-presentation condition, children were repeatedly exposed to the toy names through the experimenter's repetition and elicitation, matched in schedule and frequency to the expanding retrieval practice condition, isolating the effect of retrieval practice. In the massed elaboration condition, the presentation of the material involved evenly distributed total time with one extended discussion per toy, during which the experimenter introduced, repeated, and elicited the toy name as often as in the other conditions. Three post-practice tests were administered, one with a 1-min delay, one with a 1-day delay, and another one with a 2-day delay. The result relevant for our purposes is that the beneficial effect of retrieval practice over the two other conditions was observed for the first two post-practice tests, showing that the method of retrieval-based learning may also be useful already in this age group.

In another study Kliegl et al. (2018) addressed questions regarding methods of presentation and testing. Specifically, the authors tested the effects of retrieval practice employing free recall, cued recall, and feedback in a series of four experiments. Their main finding is that preschool children show the effect of retrieval practice, especially when cued-recall is employed during practice combined with immediate feedback, and if cued recall is also employed in the final test.

Furthermore, in a series of studies, Leonard and his colleagues investigated the retrieval practice effect in typically developing preschool children and in those with developmental language disorder (Haebig et al. 2021; Leonard, Deevy, et al. 2019; Leonard et al. 2020, 2022, 2023, 2024; Leonard, Karpicke, et al. 2019; Leonard and Deevy 2020).

Gordon and Lowry (2024) took another approach by examining how the number of sessions in which words are successfully retrieved during training affects long-term retention in 4–6-year-old preschool children. Participants undertook six training sessions containing blocks of retrieval practice, and the retrieval success of each item was registered for each session. Any item that was correctly named was dropped from the training blocks of that session. The researchers tested whether the number of sessions in which each item was correctly produced was positively related to the probability of a successful retrieval of that item on a test that was administered 1 month later. The results confirmed this relation. While previous research focused on retrieval frequency within single sessions, this study emphasizes the benefits of multiple retrieval sessions. It suggests that children achieve better long-term retention by retrieving words across several sessions rather than just once. However, a key limitation, as noted by the authors, is the relatively small sample size of 24 children.

Kueser et al. (2021) investigated how children with developmental language disorder and typically developing children learn words using repeated spaced retrieval. The authors analyzed trial-by-trial learning data from three studies, fitting growth curves to examine immediate and spaced retrieval trials. Given that these studies included breaks between sessions, it was particularly relevant to investigate the trajectory of learning. Results indicated that linear growth provided the best fit for this trajectory, suggesting that the effect of spaced retrieval is robust despite interruptions. Additionally, incorporating immediate retrieval trials during spaced retrieval was found to enhance success in spaced retrieval. Finally, retention was predicted by both cumulative success during spaced retrieval—measured as the number of successful item retrievals during study—and the overall learning trajectory.

The key takeaway from the studies surveyed in the present section for us is that they demonstrate a retrieval practice effect in age groups (from 48 to 71 months) similar to those we aim to investigate. However, it is important to note that

these studies used relatively small sample sizes ranging from 10 to 16 participants per group (except for Kliegl et al. 2018). Additionally, in general, the studies mentioned above utilized learning schedules spanning 2 weeks, with two 20-min learning sessions each week, and a majority of them involved a limited amount of material to be learned, specifically three to four words per condition (except for e.g., Gordon et al. 2021; Gordon and Lowry 2024). Therefore, we believe that this question should be addressed in a study utilizing a much larger sample size, a larger study set of six items, and a shorter learning session. The last two considerations are practical: in educational applications, our goal is to maximize efficiency. A tool is more efficient if it helps students learn more material in a shorter period of time.

2 | Research Objectives

Our research is motivated by three considerations. First, relatively few studies have investigated the effect of retrieval practice in children, and these studies have used small sample sizes and variable ranges of age. For this reason, our objective is to investigate the effect of retrieval practice on a large sample. We consider it crucial to use larger sample sizes to bolster confidence in both the presence and stability of the retrieval practice effect within this age group. Based on earlier findings (see literature survey in the preceding section), we predict that preschool children between the age of 5 and 6 also benefit from the method of retrieval practice. Specifically, we hypothesized that children who learn using retrieval practice would outperform those who learn by restudy on both final tests of recall and recognition. Second, and more importantly, few studies have investigated the role of initial learning in the emergence of the retrieval practice effect in adults, and very few studies have investigated it in preschoolers (e.g., Gordon and Lowry 2024; Leonard et al. 2020). For this reason, we aimed to address the question of whether increasing the amount of initial learning is an efficient means of improving retrieval performance during retrieval practice, and whether it may thus lead to a better long-term memory performance. Based on previous findings on adults (e.g., Racsmány et al. 2020), we suggest that the amount of initial learning modulates the emergence of the retrieval practice effect. In other words, the probability of the retrieval practice effect occurring is higher when a greater amount of initial learning precedes the practice. Third, our survey of the methodologies of earlier studies revealed that the effect of retrieval practice could be detected after lengthy training of 20-min sessions spanning 2 or more days. Hence, we aimed to investigate if it is possible to detect the effect with a shorter learning phase consisting of a single 20-min session. We consider this an important question to address, as one aim in both everyday life and educational settings might be to learn as much as possible in the shortest possible time.

3 | Materials and Methods

We conducted two experiments. As one of our main questions focused on how the amount of initial learning influences the retrieval practice effect, we structured the experiments identically, varying only the amount of initial learning. Each experiment included a pre-practice initial learning block, and then we introduced either restudy or retrieval practice as manipulated conditions. The key difference between the experiments was that we doubled the amount of initial learning by incorporating two learning blocks in the second experiment instead of one.

3.1 | Participants

A total of 102 White Hungarian middle-class preschoolers (53 girls; $M_{age} = 66.00$ months, $SD_{age} = 2.90$, range: 60–72) participated in the first experiment. In the second experiment, 100 middle-class preschoolers (43 girls; $M_{age} = 65.87$ months, $SD_{age} = 3.72$, range: 60–72) participated. Data were collected between 2019 and 2024. We consulted with the kindergarten teachers to ensure that all participants were typically developing children with no known psychological, linguistic, developmental disorders, or hearing or visual impairments.

Recruitment took place in kindergartens in Hungary between 2020 and 2024. The research was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. Prior to the experiments, we informed parents about the study, who signed a consent form for the participation of their children.

3.2 | Stimuli

The stimulus set contained six CVC nonwords and six novel animal figures that were unknown to the children. The nonwords were created based on the following considerations: we selected consonants that are easily pronounceable for typically developing children of this age, and vowels that are relatively distant from each other on the Hungarian vowel trapeze, ensuring that the contrast between the nonwords would be easily perceivable. The resulting six words were the following: *tács* /taːtʃ/, *géb* /geːb/, *mic* /mit͡s/, *löb* /løb/, *pug* /pug/, *nyüf* /ɲyf/. The animal figures were drawn by an artist. The artist was asked to draw the figures in a way that each possesses a distinctive character, yet none of them singularly stands out from the rest. This was achieved by avoiding the use of particularly vibrant colors or unusual shapes that could draw undue attention. The six animal figures are presented in Figure 1A.

In order to control for any effect resulting from associations of phonological and visual form, we created two lists, in which the six nonwords and animal figures were differently paired. The participants were randomly assigned to the two lists, ensuring equal proportions in each.

The experiments were recorded using the PsychoPy 3.0 experimental design software (Peirce et al. 2019) in order to ensure a controlled pacing of stimulus presentation. The instructions for the practice and test phases were recorded by a trained female speaker in a sound-treated room.

3.3 | Procedure

Sessions took place in quiet and separated rooms of the kindergartens. Before the sessions, all distractions (e.g., mirrors, toys)

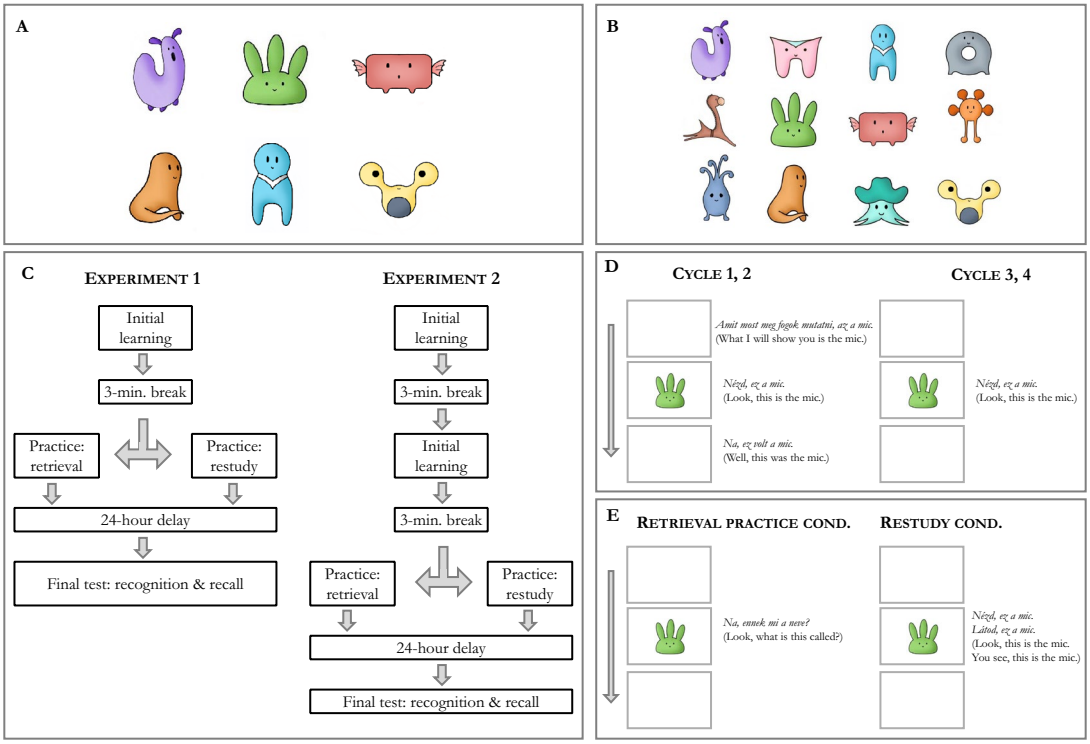

FIGURE 1 | The stimulus set and procedure of the two experiments. (A) The animal figures constituting the learning set. (B) One version of the visual stimulus used in the recognition test. (C) The procedure of the two experiments. (D) Instructions and their timing in the initial learning phase. (E) Instructions and their timing in the practice phase.

were removed. The sessions began with a brief conversation to help the children feel comfortable. Additionally, children were informed that they could choose a coloring sheet after the session. This was done to improve their motivation for engaging in the tasks. The structure of the two experiments is presented in Figure 1C.

Both experiments consisted of three phases: an initial learning phase, a practice phase, and a final test phase. The initial learning phase consisted of one learning block in the first and two learning blocks in the second experiment. This was the only difference between the two experiments. During one block, the entire study set (i.e., six name-animal pairs) was presented eight times in four cycles (3×, 3×, 1×, 1×). The presentation order of the animals within these cycles was randomized.

During the presentations of the initial learning phase, each animal figure appeared in the middle of the screen and was named by the instruction. The specific instructions and their timing relative to the visual stimulus presentation during the initial learning phase are presented in Figure 1D.

The initial learning phase was followed by a 3-min break, during which children engaged in nonverbal play, such as throwing balls or playing with toy cars. After the break, the practice phase began, consisting of four cycles. The cycles were identical throughout the practice phase within each condition. Children were randomly assigned to conditions, ensuring an equal proportion in each.

In the restudy condition, children were presented with the names of the animals two times in each block. For each study item, the animal figure appeared and the participant heard the following instruction: *Look, this is the* ... After this, the animal figure was named again: *You see, this is the* ... Subsequently, the figure disappeared, and the next presentation began.

In the retrieval practice condition, the study set was practiced using cued recall (Kliegl et al. 2018) with feedback (Butler and Roediger 2008; Kang et al. 2007). For each study item, the animal figure appeared on the screen, and the participant heard the following question: *Look! What is this called?* If the participant's response was correct, the following feedback instruction was played: *Yes, this is the* ... If the participant produced an incorrect name, the feedback named the animal figure: *This is not what it's called. This is a* ... If the participant did not produce any name, the instruction was as follows: *Look! This is the* ... Subsequently, the figure disappeared, and the next presentation began. The items of the study set were presented in a random order.

The rationale for presenting the animal names twice in the restudy condition was to closely match the exposure level in the retrieval practice condition (for similar methods, see Szőllősi et al. 2017). In this condition, the child encountered the name twice: once during the retrieval attempt (or an attempt to produce it) and again during feedback. By presenting the names twice in the restudy condition, we attempted to achieve comparable exposure across the two conditions.

The first session ended after the practice block. The entire session lasted for approximately 20 min. At the end, children could choose a coloring sheet and were informed that on the next day they would meet the experimenter again, and their task would be to remember the names of the animal figures.

The final test phase began 24 h after the first session. At the beginning, children were promised a coloring sheet again. The test phase consisted of two tasks: a recall test and an associative recognition test (recognition test, henceforth). The recall test was run on the computer. During the trials, children were presented with each animal figure in the middle of the screen and heard the following instruction: *Look! What is this called?* The order of the items was randomized also within the recall test. Next, the recognition test ensued. During the test, the experimenter placed a sheet containing all the animal figures and an equal amount of lures in a grid-like orientation in front of the participant. Several such grids were created, in which the arrangement of the figures was quasi-randomly distributed. Children were assigned to the grids randomly. One example is presented in Figure 1B.

After placing the grid in front of the child, the experimenter produced the names of each of the figures using the carrier sentence *Where is the …?*, and asked the child to show the corresponding figure on the grid. After the recognition test, the experiment ended, and the children could choose a coloring sheet. One session in the final test phase lasted approximately 5 min.

4 | Results

During the sessions, the experimenter registered responses using the computer keys. Three types of responses were registered: (i) incorrect or no response; (ii) partially correct response, in which at least the vowel of the nonword and one consonant were correct (e.g., *kic*/kiʧs/instead of *mic*/miʧs/) and (iii) correct response. Only correct responses were included in the analyses below.

As a first step, we examined whether sex or age influenced memory performance, as these factors may impact performance in this age group. The details of the analysis are presented in Supporting Information S1. In brief, the analysis of sex and age differences revealed no significant effects on either recall or recognition performance, either in the practice phase or on the final tests. We analyzed recall performance change from the first to the final cycle of the practice phase, performance on the final practice cycle, and performance on the final recall and recognition tests. Additionally, we calculated the forgetting rate as the difference between recall performance on the final test and the final cycle of the practice phase. The results indicated no statistically significant age-related relations and sex differences across any of the measures analyzed. Based on these results, we concluded that sex and age should not be entered as covariates in the statistical analyses on performance.

Subsequently, we analyzed retrieval performance during the training cycles. Figure 2A shows retrieval success during the practice cycles in the two experiments. As can be seen in the plot, retrieval success was higher in Experiment 2 than in Experiment 1. As the only difference between the two experiments was the amount of initial learning, this difference is attributable to a higher prior knowledge achieved in Experiment 2. The rates of retrieval success were analyzed using one-way ANOVA with the factor cycle in the case of both experiments. Analyses showed that retrieval success improved throughout the cycles in both experiments: $F\ (3,\ 208) = 5.644$, $p < 0.001$, $\eta^2 = 0.075$ in Experiment 1, and $F\ (3,\ 196) = 9.550$, $p < 0.001$, $\eta^2 = 0.127$ in Experiment 2.

Next, we analyzed the recall performance rates (see Figure 2B) on the final tests using Mann–Whitney *U* tests in the case of both experiments. No difference was found in recall performance between the restudy and retrieval practice conditions in Experiment 1 ($W = 1251$, $p = 0.738$, $d = 0.034$); however, it was found that recall success was significantly better in the retrieval practice condition as compared to the restudy condition in Experiment 2 ($W = 801.5$, $p = 0.007$, $d = 0.315$).

Figure 2C shows recognition performance in the final test phases of the two experiments using Mann–Whitney *U* tests. There was no difference in recognition memory performance between the restudy and retrieval practice conditions in Experiment 1 ($W = 1081$, $p = 0.133$, $d = 0.149$); however, recall success was significantly better in the retrieval practice condition as compared to the restudy condition in Experiment 2 ($W = 807$, $p = 0.001$, $d = 0.324$).

After confirming positive answers to our research questions—specifically, (i) the retrieval practice effect is present in children aged 5–6 years, and (ii) the amount of initial learning plays a significant role in the emergence of these effects—we conducted further analyses to address two key questions.

The first question concerns how the number of successful retrievals of each item (i.e., animal name) during practice relates to long-term retention. In other words, we investigated how many successful retrieval attempts were necessary for successful retrieval on the final test. We carried out this analysis for the experimental groups of both experiments. For both experiments, we first calculated the number of successful retrievals across the four practice phase cycles for each item, distinguishing between two cases: (i) when the item was successfully recalled on the final test, and (ii) when it was not. An ANOVA conducted on these rates within the successful and unsuccessful categories revealed no significant differences in the number of times each item was recalled during the training cycles. In other words, all animal names were recalled at similar rates, with none being recalled more or less frequently than the others when they were successfully recalled on the final test (Exp. 1: $F\ (5,\ 91) = 0.349$, $p = 0.882$, $\eta^2 = 0.019$, Exp. 2: $F\ (5,\ 114) = 2.068$, $p = 0.075$, $\eta^2 = 0.083$), or when they were not recalled on the final test (Exp. 1: $F\ (5,\ 225) = 0.910$, $p = 0.475$, $\eta^2 = 0.020$, Exp. 2: $F\ (5,\ 174) = 2.104$, $p = 0.067$, $\eta^2 = 0.057$). Next, we calculated the mean number of recalls during training for the two categories (see Figure 3A). If the recall of the item was successful on the final test, these values were 2.79 (SD = 0.23) in Experiment 1 and 2.77 (SD = 0.45) in Experiment 2. If the recall was unsuccessful, the mean number of recalls during practice was 0.50 (SD = 0.16) in Experiment 1 and 0.89 (SD = 0.29) in Experiment 2. A Wilcoxon rank-sum test revealed a significant difference between the two values in the case of both experiments (Exp. 1: $V = 21$, $p < 0.031$, $r = 0.899$, Exp. 2: $V = 21$, $p < 0.031$, $r = 0.899$). In brief, almost three successful retrieval attempts were necessary to recall an item on

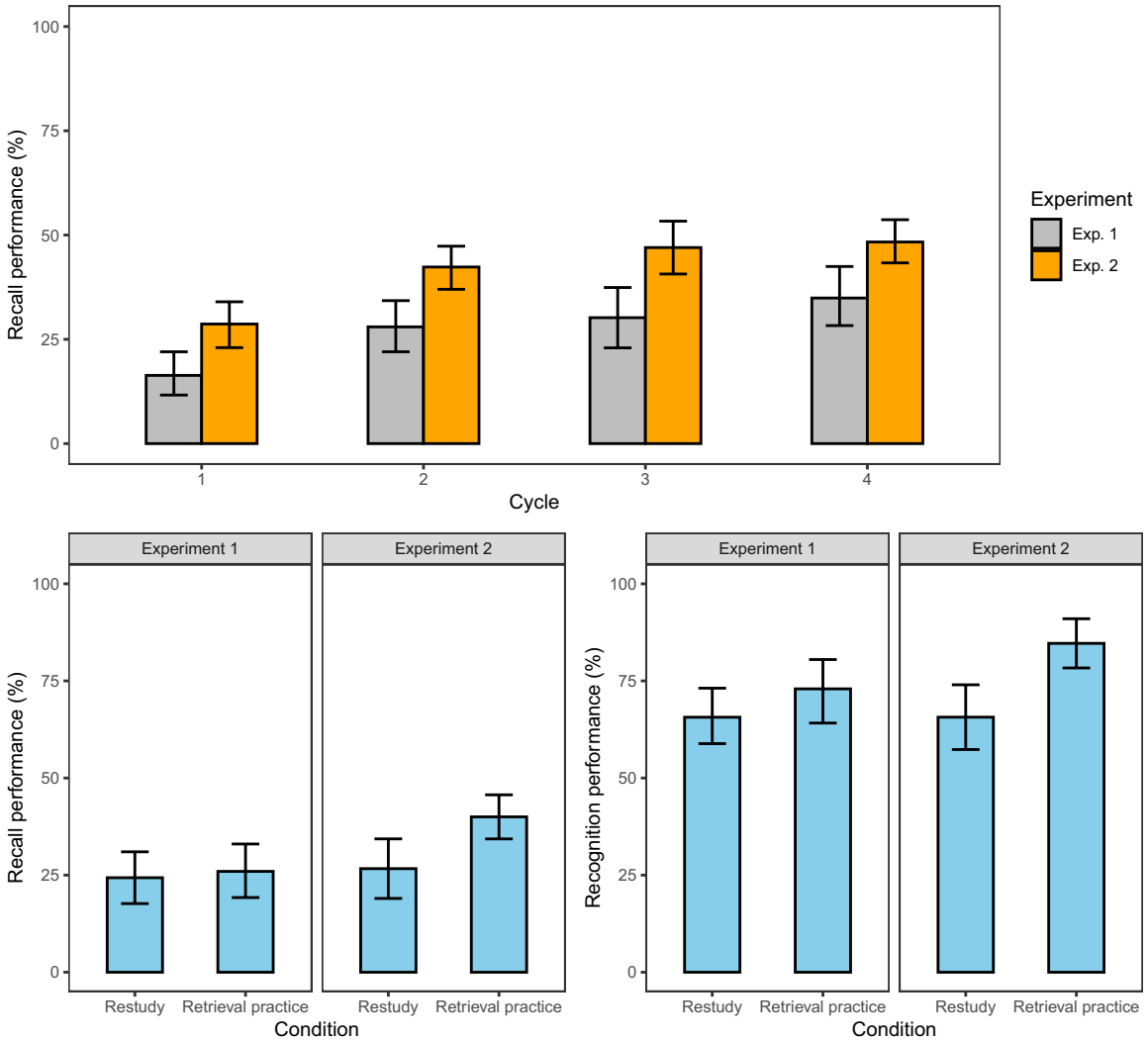

FIGURE 2 | Memory performance (error bars: 95% confidence intervals of the mean). (A) Recall performance during the practice block. (B) Recall performance on the final test. (C) Recognition performance on the final test.

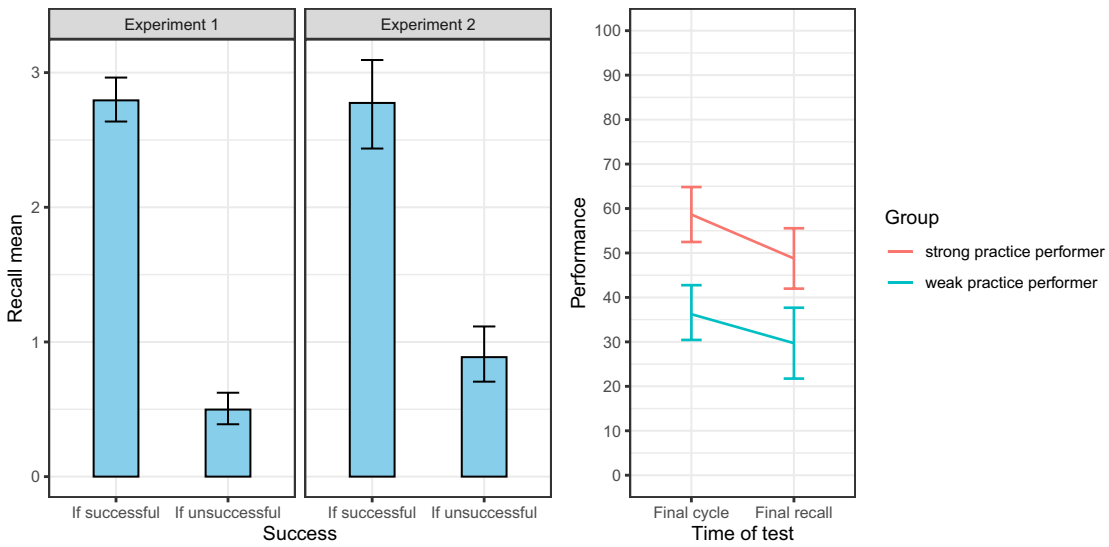

FIGURE 3 | Error bars: 95% confidence intervals of the mean. (A) Mean recall rates of animal names during training as a function of final recall success in the two experiments. (B) Performance on the final cycle of the practice phase and the final recall test of the strong and weak performer groups in Experiment 2.

the final test, supporting the assumption that multiple successful retrieval is necessary for the retrieval effect to occur.

The second question asks if those who perform better on the final cycle of the practice will perform better on the final recall test (Figure 3B). To address this question, participants were divided into two groups using a median split. Strong practice performers were those whose performance equaled or exceeded the median performance expressed as a percentage of all participants during the entire practice block (median = 41.667%), while weak practice performers were those whose performance fell below this median value. Next, we conducted a 2 (practice performance: strong vs. weak practice performers) × 2 (time of test: final cycle vs. final recall) repeated-measures mixed ANOVA. The analysis revealed a significant main effect of Practice performance (F (1, 48) = 24.02, $p < 0.001$, $\eta^2 = 0.259$), and a significant main effect of time (F (1, 48) = 8.74, $p = 0.005$, $\eta^2 = 0.052$). The practice performance × time interaction was not significant (F (1, 48) = 0.37, $p = 0.548$, $\eta^2 = 0.002$). Next, we conducted post hoc comparisons with Bonferroni correction to determine if groups differed in performance. The results revealed a significant difference between the strong practice performer group and the weak practice performer group (t (48) = 4.901, $p < 0.001$, $d = 1.143$). This suggests that participants who performed strongly during practice achieved significantly higher performance on the final recall test compared to those who performed weakly during practice. Additionally, a post hoc comparison was conducted to explore the main effect of time of test. The results showed a significant difference between the final cycle performance and the final recall performance (t (48) = 2.957, $p = 0.0048$, $d = 0.403$). This suggests that participants performed significantly better on the final cycle compared to the final recall test. One of the most important aspects of these results is the absence of a significant interaction. The lack of interaction shows that while individuals may differ in their performance during practice, the extent of these differences remains stable over the course of a single day. This suggests a general beneficial impact of retrieval practice in preschoolers.

5 | Discussion

Retrieving material during practice facilitates long-term memory retention. This beneficial effect is called the retrieval practice effect Roediger and Karpicke 2006. Additionally, there is a crucial, but often overlooked factor: retrieval during practice has to be successful for the effect to emerge. Hence, the goal of the present study was twofold: first, we aimed to investigate the retrieval practice effect in a large sample of preschoolers, an age group that has been underrepresented in prior research on this topic. We consider the large sample size and the replication of the effect especially important, as it can strengthen our confidence in the existence and robustness of the beneficial impact of retrieval practice in the given age group. Second, our goal was to investigate the role of initial learning (and hence successful retrieval) in the emergence of the retrieval effect in this particular age group. Regarding this issue, studies have predominantly focused on adults (e.g., Racsmány et al. 2020), and little attention has been given to how successful retrieval during practice sessions can influence long-term memory retention in preschool children.

5.1 | Retrieval Practice Effect and Methodological Considerations

Regarding the first question addressed in the present study, our results confirm findings in the literature by demonstrating that preschoolers do indeed exhibit the retrieval practice effect. However, it must be emphasized that the effect was only observed in the second experiment, where the amount of initial learning, and hence the retrieval success rate during practice, was higher. These results underscore the importance of retrieval success during practice and indicate that it is a necessary condition for the effect to occur. Additionally, it is also interesting to observe that while individuals may differ in their performance during practice, the extent of these differences remains stable over the course of a single day. This suggests that retrieval practice, if present, is a generally efficient method of learning in preschoolers. Thus, our findings replicate and extend those of earlier studies on this research topic. As far as replication is concerned, we consider our findings important, especially in light of the methodological diversity in studies addressing the issue.

For example, the first study on the question by Fritz et al. (2007) admittedly had some methodological shortcomings, such as an imbalanced number of stimulus presentations in the different conditions in the first experiment. The method is problematic because the level of exposure to new material significantly impacts memory retention. Nevertheless, the authors corrected this methodological issue in the second experiment. Another potential methodological weakness is the use of existing names like *Charlie* or *Tinker*, as it is possible that some children already know some of these names or even have some emotional relations to them (e.g., the family dog is called Charlie), while other names may be totally unfamiliar to them. This asymmetry may confound the memory performance of some children. However, it must be emphasized that the study by Fritz et al. (2007) is the first one to address the retrieval practice effect in preschool children and is of great importance in the literature.

An additional methodological issue is sample size. For example, sample sizes in a series of studies by Leonard and his colleagues range from 10 to 16, with a mean of 13.25 (SD = 1.98) (Haebig et al. 2021; Leonard, Deevy, et al. 2019; Leonard et al. 2020, 2022, 2023, 2024; Leonard, Karpicke, et al. 2019; Leonard and Deevy 2020). These studies investigated the retrieval practice effect not only in typically developing children, but also in children with developmental language disorder. This is an important methodological aspect, as involving group as a factor entails not only within, but also between subjects comparisons. These studies did demonstrate the presence of retrieval practice in preschool children; however, it is worth considering that the statistical power achievable with small sample sizes may be limited, potentially leading to false positives and restricted generalizability. Hence, while the results of these studies are undoubtedly intriguing and of high scientific value, we believe it is important to use a large sample size to draw firm conclusions and make generalizable claims about the population of interest. The experiments reported in the present paper achieved this aim.

Additionally, in a number of papers in which the age range of the sample is reported, it turns out that the range exceeded 12 months. For example, the age of the children ranged from 53 to 71 months (18 months) in Leonard, Deevy, et al. (2019) and Leonard, Karpicke, et al. (2019), 48 to 71 months (23 months) in Leonard et al. (2020) and from 51 to 71 months (20 months) in Leonard and Deevy (2020). Additionally, the range was from 48 to 75 months (24 months) in the experiments reported by Haebig et al. (2021) and 30 to 60 months (30 months) in the first experiment by Fritz et al. (2007). We suggest that using such a wide age range may be problematic, as the development of relevant cognitive abilities in preschool children is considerable during these periods of time (Gathercole 1998; Kohlberg 1968), which makes the sample heterogeneous. For this reason, we kept the age range within 12 months in our experiments.

Another aspect is the amount of practice and the size of the study set. A survey of earlier experiments reveals that in most of the cases, practice took place on two consecutive days, with 20-min sessions on each day (Haebig et al. 2021; Leonard, Deevy, et al. 2019; Leonard et al. 2023, 2024; Leonard, Karpicke, et al. 2019; Leonard and Deevy 2020). Additionally, in these experiments, the amount of to-be-learned material was four nonword-object pairs per condition, except in Haebig et al. (2021), where it was six, and in Gordon et al. (2021) and Gordon and Lowry (2024), where it was nine. We assume that these studies used low set sizes, as they employed a within-subject design; hence, they did not want to increase the total size above eight. In the experiments reported here, we aimed to investigate whether the retrieval practice effect persists with a single 20-min practice session and a larger set size of six items. We believe that our results are interesting for two reasons. First, we confirmed that the retrieval practice effect exists even with a lower amount of practice and a higher amount of material to be learned. Second, we showed that not only the quantity of practice matters, but also sufficiently high level of success during practice, which can be facilitated through a proper amount of initial learning. This is evidenced by the difference in results between our first and second experiments: purely increasing the amount of prior learning already led to the emergence of the retrieval practice effect.

A key conclusion of the present study is that successful retrieval during practice enhances long-term memory retention. However, research on adults also suggests that even unsuccessful retrieval attempts contribute to long-term retention (see e.g., Karpicke 2009). One explanation for this effect is that attempting to retrieve information makes learners more aware of what they have already mastered and what still requires further study. Consequently, they often gain more from the subsequent exposures after retrieval attempts than from passive review alone. In line with Karpicke (2009), we propose that this mechanism operates in adults due to their metacognitive skills, which enable them to consciously optimize their learning strategies. Another line of research (Kornell et al. 2009; Richland et al. 2009) offers alternative explanations for the long-term benefits of unsuccessful retrieval attempts during practice. These studies suggest mechanisms such as facilitating deeper processing, enhancing recall cues, or strengthening the association between cues and correct responses. It appears that there is no consensus on the precise mechanisms behind the positive effects of unsuccessful retrievals. As this issue has not been extensively studied in preschool-aged children, whose metacognitive skills are not yet fully developed, an open question remains: do younger children also benefit from unsuccessful retrieval attempts? Future research is needed to explore this question.

To summarize, the studies cited above make invaluable contributions to our understanding of the mechanisms and various aspects of the retrieval practice effect, while some of their methodological aspects could be altered. Our aim was to set up an experiment in which these potential issues are addressed by using a large sample size, a relatively narrow age group, a lower amount of practice, and a higher amount of to-be-learned material. Our results confirm, and hopefully complement, the findings of earlier studies.

5.2 | The Role of Initial Learning

A key finding of our study is that the effectiveness of retrieval practice depends on reaching a certain level of performance through a sufficient amount of initial learning before practice begins. With a relatively low amount of initial learning, the success rate of retrieval was also low in the first experiment, and as a consequence, the beneficial effects of retrieval practice could not be detected. This was true even in the case of the recognition task, a form of retrieval that requires less cognitive demand as compared to recall (Atkinson and Juola 1974; Yonelinas et al. 2022). In the second experiment, however, in which the amount of initial learning was higher, retrieval success was also higher during practice. As a consequence, a considerable retrieval practice effect was observed not only in the recall phase of the final test, but also on the recognition memory test. These findings underscore the critical role of initial learning in preschool children.

An interesting aspect of our results is that retrieval success improved throughout the cycles in both experiments. This result has two potential, mutually nonexclusive interpretations. On the one hand, the improvement shows that preschoolers can use the cue (the image of the animal in our case) with increasing efficiency to find the target (the name of the animal). On the other hand, it indicates that feedback has a beneficial effect during the process of retrieval practice, as it provides an opportunity for further study during the cycles. This final conclusion aligns with the literature on the importance of feedback (see e.g., Butler and Roediger 2008; Kang et al. 2007).

Additionally, our results align with and extend previous findings in adult populations, particularly those observed by Racsmány et al. (2020), who also demonstrated that a sufficient amount of initial learning is essential for the retrieval practice effect to emerge. However, based on Pyc and Rawson (2009) it is also important to note that the level of criterion (i.e., the number of times an item is required to be correctly recalled before dropping from practice) may not only have a lower, but also an upper limit for the retrieval practice effect to emerge. More specifically, the authors found that retrieval during practice must be optimally difficult (Bjork 1994). If the amount of initial learning exceeds a certain threshold, the level of prior knowledge will be high, resulting in minimal effort to retrieve the learning material. This,

in turn, may diminish the effectiveness of retrieval practice. Returning to our original question, this suggests that an optimal level of difficulty must be achieved to maximize the benefits of retrieval practice also in preschoolers. Future research should thus explore how to systematically determine and implement this optimal criterion level to enhance learning outcomes effectively.

In sum, our study shows that retrieval practice can be an efficient learning strategy for preschool children when implemented under optimal conditions. Perhaps the most important aspect of our findings is the emphasis on not only testing, but also on ensuring successful retrieval during practice sessions (in line with Karpicke et al. 2014; Pyc and Rawson 2009; Racsmány et al. 2020), which is achieved via a sufficient amount of initial learning. However, it must be emphasized that in educational settings, vocabulary teaching typically involves selecting a set of words for a specific time period (e.g., a week) and exposing children to this set multiple times during that period. This approach aligns with the concept of 'fast mapping, slow learning' (Bion et al. 2013; McMurray et al. 2012). To enhance ecological validity, future studies should consider incorporating this aspect into their experimental procedures. Moving forward, continued research should also explore methodologies to identify appropriate learning conditions (e.g., optimal amount of initial learning), thereby maximizing the educational benefits of retrieval practice in preschool children.

5.3 | Implications for Language Development

Our findings also have implications regarding linguistic development in preschool years. A number of studies have emphasized that active verbal communication with children enhances their linguistic development by improving listening and asking strategies (Deason 2009), social interaction (Tzouriadou and Manavopoulos 1995), and improving vocabulary (Rowe 2013). Additionally, we contend that active communication with children creates natural conditions for the retrieval practice effect to occur. Specifically, when the child wants to achieve a communicative goal, it is important for her to use linguistic expressions that are most efficient for that particular goal. This involves the retrieval of the words needed, whereby the mechanisms of long-term memory encoding observed in the current study are facilitated. Apparently, there is another ingredient, namely, a sufficient amount of initial learning: the child has to be exposed to linguistic expressions frequently enough, so that these can be retrieved later. Hence, the more communicative situations the child is exposed to, the more instances of prior learning occur, increasing the opportunities for the mechanisms discussed in the present study to operate.

6 | Limitations

One limitation of our study is the imbalanced amount of exposure to the animal names during the practice phase. In the restudy condition, the names were presented twice in each trial: once upon the presentation of the animal figure and once again when the figure disappeared. In the retrieval practice condition, the animal name was presented only once as feedback: the child either produced the name or failed to produce it, and the name was presented as feedback. Although this is indeed a methodological shortcoming, we still think that it does not compromise our results for two reasons. First, due to the nature of the phenomenon under study, it is not possible to perfectly control the amount of exposure in the two conditions. This imbalance arises because retrieval success is a dependent variable even during the practice phase. As other studies also face this methodological issue (Szőllősi et al. 2017), we contend that our results are comparable to those in the literature. Second, although exposure may have been higher in the restudy condition, memory performance was still better in the retrieval practice condition, further underscoring the robustness of the effect.

Another limitation of our study is that children received a higher overall number of training blocks in Experiment 2 (two initial learning blocks and one practice phase) than in Experiment 1 (one initial learning block and one practice phase). Despite being exposed to only one initial learning block, children's practice performance still improved from block to block during the practice block. Moreover, the recognition test results in Experiment 1 exhibited a trend in the expected direction, with performance being better (though not statistically significant) in the retrieval practice condition. This suggests that some effect may have been present also in Experiment 1, but insufficient overall exposure may have prevented its detection. A more balanced design would involve keeping the total amount of exposure constant while manipulating the timing and amount of retrieval-based practice. For example, a comparison manipulating timing could contrast exposure-exposure-test–test–test with exposure-test–test–test-exposure, while a comparison manipulating the number of retrieval opportunities could contrast exposure-test–test–test–test with exposure-exposure-test–test–test or exposure-exposure-exposure-test–test. These designs would allow for a more precise investigation of the contribution of initial learning. A similar question was explored by Leonard et al. (2020), who examined the balance between retrieval-based practice and restudy in typically developing preschool children and those with developmental language disorder. Their results indicated that word learning and retention are enhanced when children have more opportunities to retrieve words during learning (see details in the Introduction section). Nevertheless, based on the results of the control conditions in both of our experiments, we maintain that our findings remain informative. Specifically, the median of recall performance expressed as a percentage was 16.7% (mean = 24.3, SD = 24.6) in Experiment 1, identical to that in Experiment 2 (mean = 26.7, SD = 28.2). Similarly, median recognition performance was 66.7% (mean = 65.6, SD = 26.7) in Experiment 1 and also 66.7% (mean = 65.7, SD = 30.8) in Experiment 2. The comparable performance rates across the control conditions of the experiments suggest that the amount of initial exposure may not significantly impact overall learning.

7 | General Conclusion

Our research provides significant findings that are pivotal for everyday learning and educational methodologies, particularly in preschool-aged children. Preschool years are crucial for acquiring numerous new pieces of information, and by providing

effective educational tools, we can greatly assist in this process. Our results have two consequences regarding the educational tools that may be used in this age group. First of all, we have demonstrated that the effect of retrieval practice is present in preschool children using a sample size that enables us to draw firm conclusions regarding the phenomenon. For this reason, we can confidently formulate the claim that the act of retrieving to-be-learned material is a useful method already in this age group. Second, we have shown that this is only a useful method, provided that retrieval is successful, which can be ensured by a sufficient amount of prior learning. Therefore, we propose that education professionals should not only use practice methods that prompt children to retrieve to-be-learned material, but should also ensure that new information is presented multiple times before transitioning to retrieval-based practices. Achieving a high level of knowledge is crucial, considering the amount of initial learning we employed for the retrieval success to be sufficient during the practice block.

Acknowledgments

The authors are grateful to Réka Vágner for her invaluable assistance in organizing data collection.

Data Availability Statement

The materials necessary to attempt to replicate the findings presented here are not publicly accessible. The analyses presented here were not preregistered. Data and analytic code is available at the following URL: https://osf.io/7nzjv/?view_only=6e2d76bcdccd46dc86a0259eea66928a.

References

Atkinson, R. C., and J. F. Juola. 1974. *Search and Decision Processes in Recognition Memory*. W. H. Freeman.

Balota, D. A., J. M. Duchek, and J. M. Logan. 2011. "Is Expanded Retrieval Practice a Superior Form of Spaced Retrieval? A Critical Review of the Extant Literature." In *The Foundations of Remembering*, 83–106. Psychology Press.

Bion, R. A. H., A. Borovsky, and A. Fernald. 2013. "Fast Mapping, Slow Learning: Disambiguation of Novel Word–Object Mappings in Relation to Vocabulary Learning at 18, 24, and 30 Months." *Cognition* 126, no. 1: 39–53. https://doi.org/10.1016/j.cognition.2012.08.008.

Bjork, R. A. 1994. "Memory and Metamemory Considerations in the Training of Human Beings." In *Metacognition: Knowing About Knowing*, edited by J. Metcalfe and A. P. Shimamura, 185–205. MIT Press.

Butler, A. C., and H. L. I. Roediger. 2008. "Feedback Enhances the Positive Effects and Reduces the Negative Effects of Multiple-Choice Testing." *Memory & Cognition* 36, no. 3: 604–616.

Deason, D. K. 2009. "Let's Talk: The Importance of Conversations With Preschoolers." *NHSA Dialog* 12, no. 4: 374–377. https://doi.org/10.1080/15240750903221164.

Dunlosky, J., K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham. 2013. "Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology." *Psychological Science in the Public Interest* 14, no. 1: 4–58.

Fazio, L. K., and E. J. Marsh. 2019. "Retrieval-Based Learning in Children." *Current Directions in Psychological Science* 28, no. 2: 111–116. https://doi.org/10.1177/0963721418806673.

Fritz, C. O., P. E. Morris, D. Nolan, and J. Singleton. 2007. "Expanding Retrieval Practice: An Effective Aid to Preschool Children's Learning." *Quarterly Journal of Experimental Psychology* 60, no. 7: 991–1004. https://doi.org/10.1080/17470210600823595.

Gathercole, S. E. 1998. "The Development of Memory." *Journal of Child Psychology and Psychiatry* 39, no. 1: 3–27.

Gordon, K. R., and S. L. Lowry. 2024. "Fostering Retention of Word Learning: The Number of Training Sessions Children Retrieve Words Positively Relates to Post-Training Retention." *Journal of Child Language* 51, no. 3: 710–719. https://doi.org/10.1017/S0305000923000752.

Gordon, K. R., H. L. Storkel, S. L. Lowry, and N. B. Ohlmann. 2021. "Word Learning by Preschool-Age Children With Developmental Language Disorder: Impaired Encoding and Robust Consolidation During Slow Mapping." *Journal of Speech, Language, and Hearing Research: JSLHR* 64, no. 11: 4250–4270. https://doi.org/10.1044/2021_JSLHR-21-00046.

Haebig, E., L. B. Leonard, P. Deevy, J. Schumaker, J. D. Karpicke, and C. Weber. 2021. "The Neural Underpinnings of Processing Newly Taught Semantic Information: The Role of Retrieval Practice." *Journal of Speech, Language, and Hearing Research* 64, no. 8: 3195–3211. https://doi.org/10.1044/2021_JSLHR-20-00485.

Hogan, R. M., and W. Kintsch. 1971. "Differential Effects of Study and Test Trials on Long-Term Recognition and Recall." *Journal of Verbal Learning and Verbal Behavior* 10, no. 5: 562–567. https://doi.org/10.1016/S0022-5371(71)80029-4.

Kang, S. H. K., K. B. McDermott, and H. L. I. Roediger. 2007. "Test Format and Corrective Feedback Modify the Effect of Testing on Long-Term Retention." *European Journal of Cognitive Psychology* 19, no. 4–5: 528–558.

Kang, S. H. K., H. Pashler, N. J. Cepeda, D. Rohrer, S. K. Carpenter, and M. C. Mozer. 2011. "Does Incorrect Guessing Impair Fact Learning?" *Journal of Educational Psychology* 103, no. 1: 48–59. https://doi.org/10.1037/a0021977.

Karpicke, J. D. 2009. "Metacognitive Control and Strategy Selection: Deciding to Practice Retrieval During Learning." *Journal of Experimental Psychology: General* 138, no. 4: 469–486. https://doi.org/10.1037/a0017341.

Karpicke, J. D. 2017. "Retrieval-Based Learning: A Decade of Progress." In *Learning and Memory: A Comprehensive Reference*, 487–514. Elsevier. https://doi.org/10.1016/B978-0-12-809324-5.21055-9.

Karpicke, J. D., M. Lehman, and W. R. Aue. 2014. "Retrieval-Based Learning: An Episodic Context Account." In *Psychology of Learning and Motivation*, vol. 61, 237–284. Elsevier. https://doi.org/10.1016/B978-0-12-800283-4.00007-1.

Karpicke, J. D., and H. L. Roediger. 2007a. "Expanding Retrieval Practice Promotes Short-Term Retention, but Equally Spaced Retrieval Enhances Long-Term Retention." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33, no. 4: 704–719. https://doi.org/10.1037/0278-7393.33.4.704.

Karpicke, J. D., and H. L. I. Roediger. 2007b. "Repeated Retrieval During Learning Is the Key to Long-Term Retention." *Journal of Memory and Language* 57, no. 2: 151–162. https://doi.org/10.1016/j.jml.2006.09.004.

Karpicke, J. D., and H. L. I. Roediger. 2008. "The Critical Importance of Retrieval for Learning." *Science (New York, N.Y.)* 319, no. 5865: 966–968. https://doi.org/10.1126/science.1152408.

Kliegl, O., M. Abel, and K.-H. T. Bäuml. 2018. "A (Preliminary) Recipe for Obtaining a Testing Effect in Preschool Children: Two Critical Ingredients." *Frontiers in Psychology* 9: 1446. https://doi.org/10.3389/fpsyg.2018.01446.

Kohlberg, L. 1968. "Early Education: A Cognitive-Developmental View." *Child Development* 39: 1013–1062.

Kornell, N., M. J. Hays, and R. A. Bjork. 2009. "Unsuccessful Retrieval Attempts Enhance Subsequent Learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35, no. 4: 989–998. https://doi.org/10.1037/a0015729.

Kueser, J. B., L. B. Leonard, P. Deevy, E. Haebig, and J. D. Karpicke. 2021. "Word-Learning Trajectories Influence Long-Term Recall in Children With Developmental Language Disorder and Typical Development." *Journal of Communication Disorders* 94: 106160. https://doi.org/10.1016/j.jcomdis.2021.106160.

Leonard, L. B., S. L. Christ, P. Deevy, J. Karpicke, and J. B. Kueser. 2024. "Retrieval Practice and Word Learning by Children With Developmental Language Disorder: Does Expanding Retrieval Provide Additional Benefit?" *Journal of Speech, Language, and Hearing Research* 67, no. 5: 1530–1547. https://doi.org/10.1044/2024_JSLHR-23-00528.

Leonard, L. B., and P. Deevy. 2020. "Retrieval Practice and Word Learning in Children With Specific Language Impairment and Their Typically Developing Peers." *Journal of Speech, Language, and Hearing Research* 63, no. 10: 3252–3262. https://doi.org/10.1044/2020_JSLHR-20-00006.

Leonard, L. B., P. Deevy, S. Horvath, S. L. Christ, J. Karpicke, and J. B. Kueser. 2023. "Can Retrieval Practice Facilitate Verb Learning in Children With Developmental Language Disorder and Their Peers With Typical Language Development?" *Journal of Speech, Language, and Hearing Research* 66, no. 4: 1309–1333. https://doi.org/10.1044/2022_JSLHR-22-00509.

Leonard, L. B., P. Deevy, J. D. Karpicke, et al. 2019. "Adjective Learning in Young Typically Developing Children and Children With Developmental Language Disorder: A Retrieval-Based Approach." *Journal of Speech, Language, and Hearing Research* 62, no. 12: 4433–4449. https://doi.org/10.1044/2019_JSLHR-L-19-0221.

Leonard, L. B., P. Deevy, J. D. Karpicke, S. L. Christ, and J. B. Kueser. 2020. "After Initial Retrieval Practice, More Retrieval Produces Better Retention Than More Study in the Word Learning of Children With Developmental Language Disorder." *Journal of Speech, Language, and Hearing Research* 63, no. 8: 2763–2776. https://doi.org/10.1044/2020_JSLHR-20-00105.

Leonard, L. B., J. Karpicke, P. Deevy, et al. 2019. "Retrieval-Based Word Learning in Young Typically Developing Children and Children With Developmental Language Disorder I: The Benefits of Repeated Retrieval." *Journal of Speech, Language, and Hearing Research* 62, no. 4: 932–943. https://doi.org/10.1044/2018_JSLHR-L-18-0070.

Leonard, L. B., J. B. Kueser, P. Deevy, E. Haebig, J. D. Karpicke, and C. Weber. 2022. "The Contributions of Immediate Retrieval and Spaced Retrieval to Word Learning in Preschoolers With Developmental Language Disorder." *Autism & Developmental Language Impairments* 7: 239694152210776. https://doi.org/10.1177/23969415221077652.

McMurray, B., J. S. Horst, and L. K. Samuelson. 2012. "Word Learning Emerges From the Interaction of Online Referent Selection and Slow Associative Learning." *Psychological Review* 119, no. 4: 831–877. https://doi.org/10.1037/a0029872.

Newcombe, N., S. L. Benear, C. Ngo, and I. R. Olson. 2022. "Memory in Infancy and Childhood." Preprint s.

Pashler, H., N. J. Cepeda, J. T. Wixted, and D. Rohrer. 2005. "When Does Feedback Facilitate Learning of Words?" *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31, no. 1: 3–8. https://doi.org/10.1037/0278-7393.31.1.3.

Pashler, H., D. Rohrer, N. J. Cepeda, and S. K. Carpenter. 2007. "Enhancing Learning and Retarding Forgetting: Choices and Consequences." *Psychonomic Bulletin & Review* 14, no. 2: 187–193. https://doi.org/10.3758/BF03194050.

Pastötter, B., and K.-H. T. Bäuml. 2016. "Reversing the Testing Effect by Feedback: Behavioral and Electrophysiological Evidence." *Cognitive, Affective, & Behavioral Neuroscience* 16: 473–488.

Peirce, J., J. R. Gray, S. Simpson, et al. 2019. "PsychoPy2: Experiments in Behavior Made Easy." *Behavior Research Methods* 51, no. 1: 195–203. https://doi.org/10.3758/s13428-018-01193-y.

Pyc, M. A., and K. A. Rawson. 2009. "Testing the Retrieval Effort Hypothesis: Does Greater Difficulty Correctly Recalling Information Lead to Higher Levels of Memory?" *Journal of Memory and Language* 60, no. 4: 437–447. https://doi.org/10.1016/j.jml.2009.01.004.

Racsmány, M., Á. Szőllősi, and M. Marián. 2020. "Reversing the Testing Effect by Feedback Is a Matter of Performance Criterion at Practice." *Memory & Cognition* 48, no. 7: 1161–1170. https://doi.org/10.3758/s13421-020-01041-5.

Richland, L. E., N. Kornell, and L. S. Kao. 2009. "The Pretesting Effect: Do Unsuccessful Retrieval Attempts Enhance Learning?" *Journal of Experimental Psychology: Applied* 15, no. 3: 243–257. https://doi.org/10.1037/a0016496.

Roediger, H. L. I., and J. D. Karpicke. 2006. "The Power of Testing Memory: Basic Research and Implications for Educational Practice." *Perspectives on Psychological Science* 1, no. 3: 181–210.

Roediger, H. L. I., A. L. Putnam, and M. A. Smith. 2011. "Ten Benefits of Testing and Their Applications to Educational Practice." In *Psychology of Learning and Motivation*, edited by J. P. Mestre and B. H. Ross, vol. 55, 1–36. Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00001-6.

Rowe, M. L. 2013. "Decontextualized Language Input and Preschoolers' Vocabulary Development." *Seminars in Speech and Language* 34, no. 4: 260–266.

Rowland, C. A. 2014. "The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect." *Psychological Bulletin* 140, no. 6: 1432–1463. https://doi.org/10.1037/a0037559.

Storm, B. C., M. C. Friedman, K. Murayama, and R. A. Bjork. 2014. "On the Transfer of Prior Tests or Study Events to Subsequent Study." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40, no. 1: 115.

Szőllősi, Á., A. Keresztes, B. Novák, B. Szászi, S. Kéri, and M. Racsmány. 2017. "The Testing Effect Is Preserved in Stressful Final Testing Environment." *Applied Cognitive Psychology* 31, no. 6: 615–622. https://doi.org/10.1002/acp.3363.

Tzouriadou, M., and K. Manavopoulos. 1995. "Preschoolers Linguistic-Cognitive Abilities and Verbal Production and Interaction." *European Early Childhood Education Research Journal* 3, no. 1: 47–60. https://doi.org/10.1080/13502939585207681.

Wheeler, M. A., and H. L. I. Roediger. 1992. "Disparate Effects of Repeated Testing: Reconciling Ballard's (1913) and Bartlett's (1932) Results." *Psychological Science* 3, no. 4: 240–246.

Yonelinas, A. P., M. M. Ramey, C. Riddell, M. Kahana, and A. Wagner. 2022. "Recognition Memory: The Role of Recollection and Familiarity." In *The Oxford Handbook of Human Memory*. Oxford University Press.

Supporting Information

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