



OPEN ACCESS

EDITED BY
Luciano Grüttnner Buratto,
University of Brasilia, Brazil

REVIEWED BY
Laurence B. Leonard,
Purdue University, United States
Simona Ghetti,
University of California, Davis, United States

*CORRESPONDENCE
Melina L. Knabe
✉ mknabe@wisc.edu

RECEIVED 01 August 2023
ACCEPTED 16 November 2023
PUBLISHED 15 December 2023

CITATION
Knabe ML, Bredemann CA and Vlach HA (2023)
A lack of retrieval practice effects in children's
recognition and recall memory of novel words
during storybook reading.
Front. Dev. Psychol. 1:1270938.
doi: 10.3389/fdyps.2023.1270938

COPYRIGHT
© 2023 Knabe, Bredemann and Vlach. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

A lack of retrieval practice effects in children's recognition and recall memory of novel words during storybook reading

Melina L. Knabe*, Catherine A. Bredemann and Haley A. Vlach

Department of Educational Psychology, University of Wisconsin-Madison, Madison, WI, United States

Introduction: Retrieval practice enhances adults' long-term memory over restudying. However, it is less clear whether young children also benefit from retrieval practice. This study assessed whether retrieval practice could support children's novel word learning in the context of storybook reading.

Method: In Experiments 1–3, preschoolers ($N = 167$, $\text{Range}_{\text{Age}} = 26\text{--}68$ months) were presented with 10 novel word-object pairs across four consecutive storybook readings. Half of the children were read the storybook four times (i.e., SSSS), whereas the other half of the children were read the storybook once and were then asked to retrieve the novel words during the subsequent three readings (i.e., STTT). Children's recall and recognition memory for the novel words was tested after a 5-min delay.

Results and discussion: Results revealed that although children had high retrieval success during learning, retrieval practice did not provide an advantage over re-reading for children's recall (Experiment 1), even when additional mapping and retrieval opportunities were provided (Experiment 2); that is, children who engaged in retrieval practice did not outperform children who re-read the storybooks. A retrieval practice effect also did not emerge in children's recognition memory of novel words (Experiment 3). Taken together, this study suggests that retrieval practice may only benefit young children under specific learning conditions. Future research should therefore consider characteristics of the learner and broader learning context to explain developmental differences in retrieval practice effects.

KEYWORDS

retrieval practice, language development, word learning, storybook reading, testing effect

Introduction

Words serve as the building blocks of language. Learning words also has important implications for later learning; children's vocabulary size predicts an array of academic and life outcomes (e.g., [Bleses et al., 2016](#); [Morgan et al., 2016](#); [Masrai and Milton, 2018](#)). Because language ability is critical for children's development, researchers across disciplines have studied the conditions that promote word learning. Specifically, researchers have been interested in identifying how children remember words across time ([Horst and Samuelson, 2008](#); [Kucker and Samuelson, 2012](#); [Vlach, 2019](#)). This work has revealed that characteristics of the learning environment can influence the degree to which children remember words (e.g., [Vlach et al., 2008](#); [Vlach and Sandhofer, 2012](#); [Twomey et al., 2014](#)).

One well-studied condition of the learning environment is *retrieval practice* (also called the testing effect; Roediger and Karpicke, 2006; Karpicke and Roediger, 2007; Pashler et al., 2007; Kornell et al., 2009; Roediger and Butler, 2011; Marsh et al., 2012). Studies on retrieval practice have found that practice testing oneself on information is more effective for learning than simply re-studying the information. In a typical retrieval practice paradigm (e.g., Karpicke and Roediger, 2007; Karpicke and Blunt, 2011), participants are assigned to one of two conditions: In the first condition, participants study (S) the information, such as a word list, before testing (T) themselves repeatedly (i.e., STTT condition). In the second condition, participants study the information repeatedly and do not test themselves (i.e., SSSS condition). Learners are then asked to recall the learned information (e.g., words) at a delayed post-test. Results reveal that participants who engage in retrieval practice outperform participants who only engage in restudying (McDaniel et al., 2007; Rowland, 2014; Karpicke and Aue, 2015).

This learning strategy has promoted long-term memory across several contexts. For instance, retrieval practice enhances memory for information in children and adults (e.g., Roediger and Karpicke, 2006; Roediger et al., 2011; Marsh et al., 2012; Coane, 2013; Maddox and Balota, 2015). Retrieval practice is also effective for learning different types of information, such as simple word lists (Bouwmeester and Verkoeijen, 2011; Goossens et al., 2014; Karpicke et al., 2016) and more educationally-relevant information, such as geography knowledge, encyclopedic facts, and science concepts (e.g., Rohrer et al., 2010; McDaniel et al., 2013; Jaeger et al., 2015). Indeed, researchers have identified the testing effect as one of the most replicable learning phenomena in psychological science (e.g., Dunlosky et al., 2013; Rowland, 2014).

Several theories have been proposed to explain the retrieval practice effect (for reviews, see Rowland, 2014; Endres and Renkl, 2015), attributing the effect to general retrieval and forgetting dynamics (e.g., elaborative retrieval theory; Carpenter, 2009). These theories posit that learners begin to forget information shortly after studying, which makes retrieval—or the process of obtaining information from memory—more effortful because the information becomes less accessible to the learner. Counterintuitively, learners benefit from effortful retrieval. The act of retrieving slows forgetting of information and increases the likelihood of retrieving the learned information after a time delay (Bjork, 1988; Bahrick and Hall, 2005; Pyc and Rawson, 2009). As a result, learners who practice testing themselves outperform learners who do not, leading to the observed retrieval practice effect.

Although retrieval practice appears to be a robust learning strategy across the lifespan, there is a clear gap in the literature: studies on retrieval practice effects in children largely focus on those who have started formal schooling (i.e., children who are 6 years or older). Thus, we do not know much about the conditions under which retrieval practice can support learning in 0- to 5-year-old children (see Fazio and Marsh, 2019, for a discussion of this issue). Existing studies on retrieval practice in young children have, however, rendered mixed results. For instance, retrieval practice, when embedded in a spaced schedule, enhanced preschoolers' memory for words and concepts (e.g., Fritz et al., 2007; Ritchie

et al., 2013). This finding was also observed in preschoolers' memory for past events (Hudson, 1990; Hudson and Sheffield, 1998). However, other studies have observed no benefit of retrieval practice in young children (Poole and White, 1991; Ornstein et al., 2006; Goossens et al., 2014; Karpicke et al., 2014). For instance, preschoolers performed comparably on a cued-recall task when the experimenter provided the word for a cue (snow-white) vs. when children were asked to generate the word themselves (e.g., snow-____; Carneiro et al., 2018). What might explain these mixed findings?

One explanation is that children's memory abilities vary considerably and thus moderate the magnitude of the retrieval practice effect (Bouwmeester and Verkoeijen, 2011). For instance, prior work has shown that young children rapidly forget information and thus struggle to retrieve previously learned information (e.g., Brainerd et al., 1990; Ghetti and Angelini, 2008; Vlach and Sandhofer, 2012). This might be a result of information being poorly learned initially (low storage strength) and/or an inability to retrieve learned information well (low retrieval strength) (Bjork and Bjork, 1992). Retrieval practice may therefore be too taxing to benefit young children's word learning; that is, children forget the learned information too quickly and fail to retrieve the information after a delay, leading to the absence of a retrieval practice effect. Indeed, prior studies have shown that additional retrieval supports, such as partially completed answer sheets, are needed to elicit a retrieval practice effect in children (e.g., Karpicke et al., 2014). It is therefore plausible that young learners only benefit from retrieval practice when appropriate supports—or scaffolds—are provided during learning.

Current study

The present study addressed the lack of research on retrieval practice in young children by assessing whether retrieval practice can promote memory for novel words learned from storybooks. Storybook reading was chosen as the experimental context for the study because it is a common word learning medium for young children in experimental and naturalistic studies (e.g., Montag et al., 2015; Flack et al., 2018). Storybooks expose children to new words and serve as important word learning opportunities for preschoolers (e.g., Silverman, 2007; Wasik et al., 2016).

Experiment 1 examined whether retrieval practice enhanced children's recall memory of novel words learned via storybook reading. Preschool-aged children were read a storybook consecutively four times where they learned 10 novel words. Half of the children heard the novel words read on all four storybook readings (i.e., SSSS), whereas the other half of the children heard the novel words read on the first reading and were then asked to retrieve the newly learned words on the next three readings (i.e., STTT). This paradigm models what has been used in previous studies of retrieval practice (Karpicke and Roediger, 2007; Karpicke and Blunt, 2011). Children's retrieval memory for the new words was assessed after a 5-min delay. Experiment 2 assessed whether a retrieval practice effect could be elicited if additional word mapping and retrieval opportunities were provided during the storybook readings. Finally, Experiment 3 had a similar

design to Experiments 1–2, except children were tested on their recognition memory of the novel words after a 5-min delay.

We hypothesized that retrieval practice would benefit children's memory for newly learned words. The retrieval practice effect is a highly replicable phenomenon and evidence for the benefit of retrieval practice has been found in children (e.g., Hudson, 1990; Hudson and Sheffield, 1998; Fritz et al., 2007; Ritchie et al., 2013). The alternative hypothesis was that children would not benefit from retrieval practice. Indeed, previous research has sometimes observed null effects, which have been attributed to significant memory constraints in early childhood that impact how well information is stored and retrieved (Poole and White, 1991; Ornstein et al., 2006; Goossens et al., 2014; Karpicke et al., 2014). Taken together, these studies were designed to expand our understanding of the conditions under which retrieval practice is beneficial for children's memory for words.

Experiment 1

Experiment 1 assessed whether retrieval practice would support children's recall memory for words. Across two conditions, children engaged in passive word mapping (i.e., studying) or retrieval practice (i.e., testing). Children's retrieval success was measured during learning. After a delay, children's memory was assessed using a recall memory task where they were asked to produce the word for a referent. We predicted that children who engaged in retrieval practice during storybook reading would outperform children who engaged in passive word mapping.

Method

Participants

The participants were 65 typically developing preschool-aged children ($M_{\text{age}} = 45.20$, $SD = 9.53$; range: 26–67 months, 34 girls) recruited from local daycare centers and preschools. An additional two children participated in the study but were not included in the final sample because of fussiness and inability to follow directions during the experiment. This age range was chosen because it is a period of early development in which children are actively acquiring new words. A power analysis, using Cohen's $s d$ and an estimate of a large effect size ($d = 0.85$), was conducted to determine a sample size that will provide at least 80% power. An effect size was collected from a retrieval practice effect study that was closest to the age range in our target sample ($ds > 1.0$; e.g., Fritz et al., 2007). A power analysis (GPower 3.1.) for an independent samples t -test, with $\alpha = 0.05$, yielded a sample size of at least 25 participants per condition to achieve 90% power. Thus, we collected data for 6 months or until we reached at least 50 participants. The participants came from a convenience sample or primarily middle- to upper-socioeconomic status households. Children received a storybook as a thank you for their participation.

Apparatus and stimuli

The experimental stimuli (Figure 1) were presented in a storybook format with 10 laminated pages ($9'' \times 11''$). The

storybook format was chosen for the word learning task because research demonstrates that children's storybooks can be facilitative for word learning and imitate a naturalistic word learning setting (e.g., Ninio, 1983). The storybook narrative ("My Very Own Birthday Cake") followed a child named Davey who bakes a cake for his birthday. Visual stimuli consisted of two-dimensional colorized line drawings of scenes, characters, familiar objects (e.g., flower, cookie jar, dog, carrots, milk, mixing bowl, spoon, candles, birthday cake) and novel objects. Scene settings included a bedroom, a kitchen, the front of a house, a grocery store, and a garden. These scenes were selected because they are all familiar spaces for children. Each scene either displayed the main character ("Davey"), or the main character with his friend ("Shannon"), baby brother ("Max"), or grandmother. The novel objects were displayed on five out of the 10 pages in the presence of novel and familiar distractor objects. In total, children saw 10 novel objects that were randomly assigned a set of 10 novel labels (i.e., *wug*, *dax*, *fep*, *nola*, *blicket*, *gips*, *motch*, *boam*, *toma*, *hap*) following American English phonotactics.

Design

Included in the design were two experimental conditions which modeled other retrieval practice effect study conditions. In seminal studies of the retrieval practice effect (e.g., Roediger and Karpicke, 2006), to-be-learned items are presented to participants on either a massed schedule, in which items are presented passively (i.e., SSSS), or on a spaced schedule, in which participants engage in active retrieval of items during learning (i.e., STTT). Half of the children were assigned to a Passive Reading Condition (i.e., Read Condition, $N = 33$); half of the children were assigned to a Retrieval Practice Condition (i.e., Retrieval Condition, $N = 32$). Testing occurred after a 5-min delay so that children had to retrieve words from long-term memory.

Procedure

The experiment contained three phases: a learning phase, retention phase, and a testing phase (Figure 2). Before beginning the experiment, the children were told that they would be read a story about a boy named Davey making a birthday cake. They were also told that they would be hearing the story four times before answering questions about the story.

Learning phase

The experimenter pointed to and labeled the objects on readings 1–4 in the Read Condition (i.e., "I'll put the *wug* on the cake. This one is the *wug*."). For the Retrieval Condition, the experimenter pointed to and labeled the object on the first read-through (i.e., "I'll put the *wug* on the cake. This is one the *wug*."). Importantly, on readings 2–4, the experimenter incorporated a retrieval test (i.e., "I'll put the *wug* on the cake. What is this one called?"). Children were asked to produce the target label. No feedback was provided and children's responses to the retrieval test during storybook reading were recorded by the experimenter.

Retention phase

The learning phase was followed by a 5-min delay, during which children participated in an unrelated activity

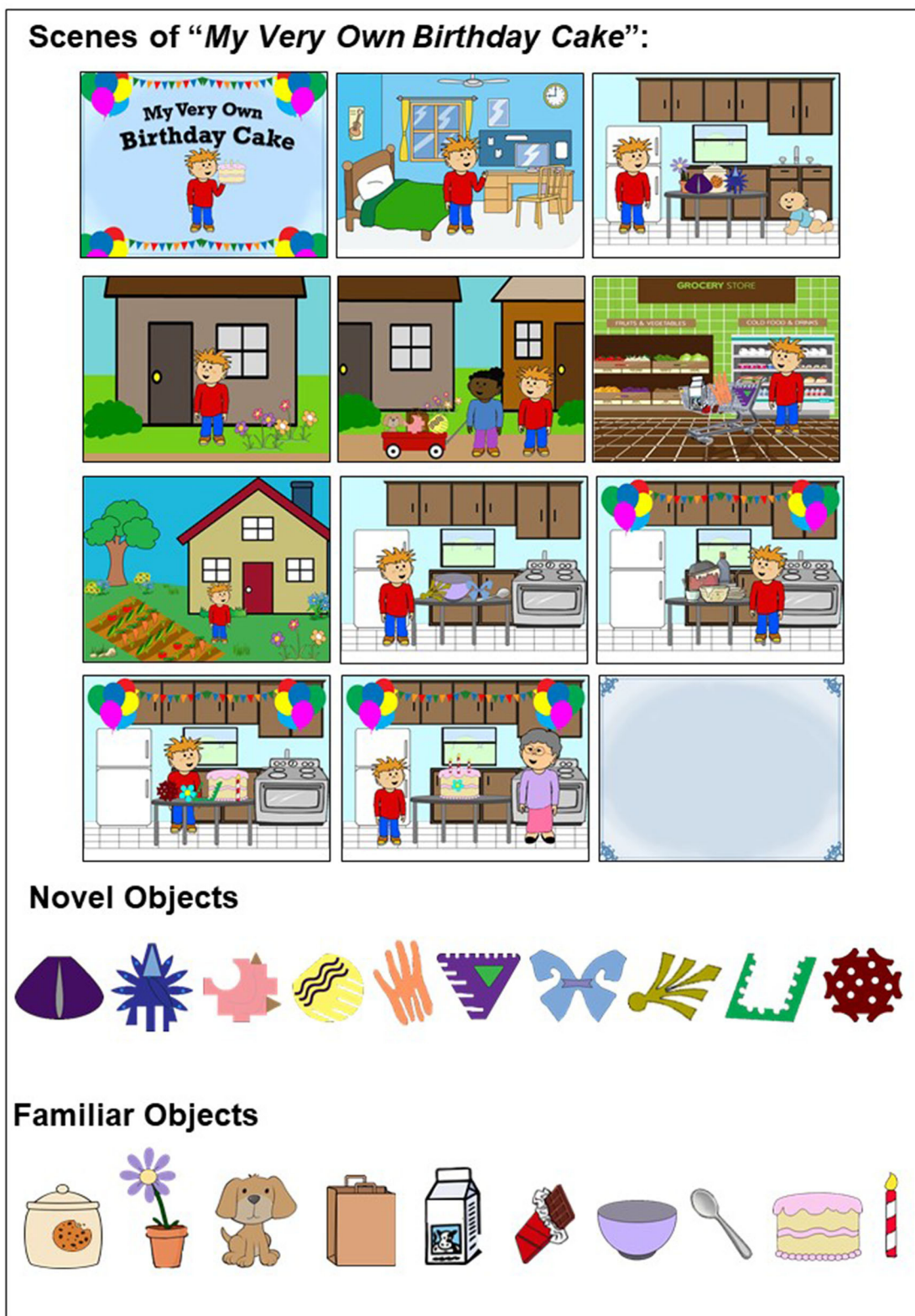
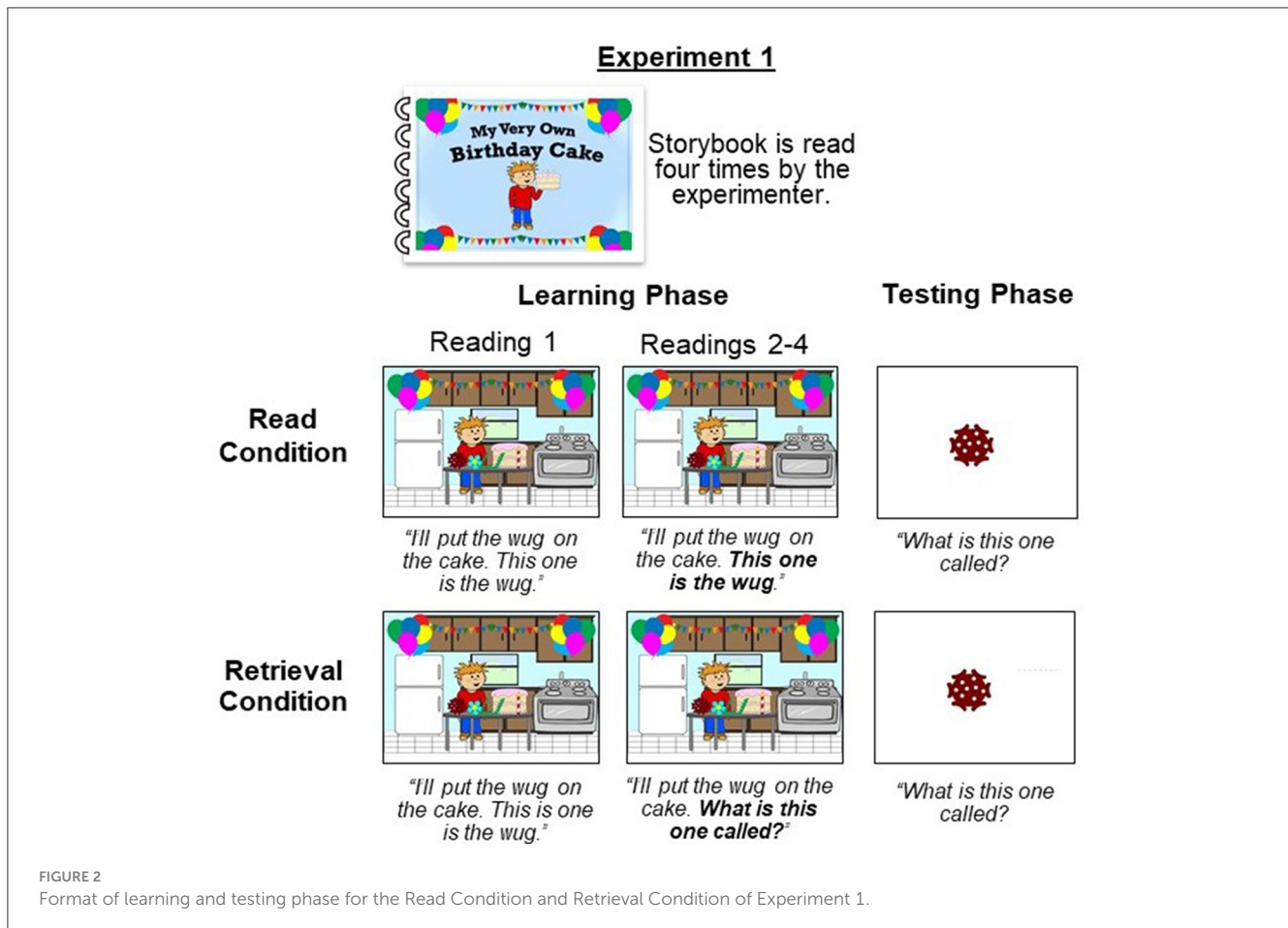


FIGURE 1 Stimuli used during the experiment. Stimuli included 10 storybook scenes, 10 novel objects, and 10 familiar objects.

(e.g., playing with a familiar toy, putting stickers on paper). This retention interval was chosen because children rapidly forget words in as short as 5 min, after which they must

retrieve information from long-term memory (Horst and Samuelson, 2008; Kucker and Samuelson, 2012; Vlach, 2019). The benefit of retrieval practice typically emerges



after a time delay when learners engage in long-term memory retrieval.

Testing phase

Following the 5-min retention interval, the experimenter presented children with a recall test consisting of 10 trials. Each trial displayed one of the target objects. The experimenter prompted the child to produce the target label by asking, "What is this one called?". Each target novel object was tested once. The experimenter did not provide feedback throughout testing. Children's answers were recorded on a response sheet as correct or incorrect. Children's responses were also coded for phonological overlap with the target label (e.g., the onset phoneme of the child's response matched the onset phoneme of the target novel label). If no answer was selected, the trial was left blank and scored as incorrect.

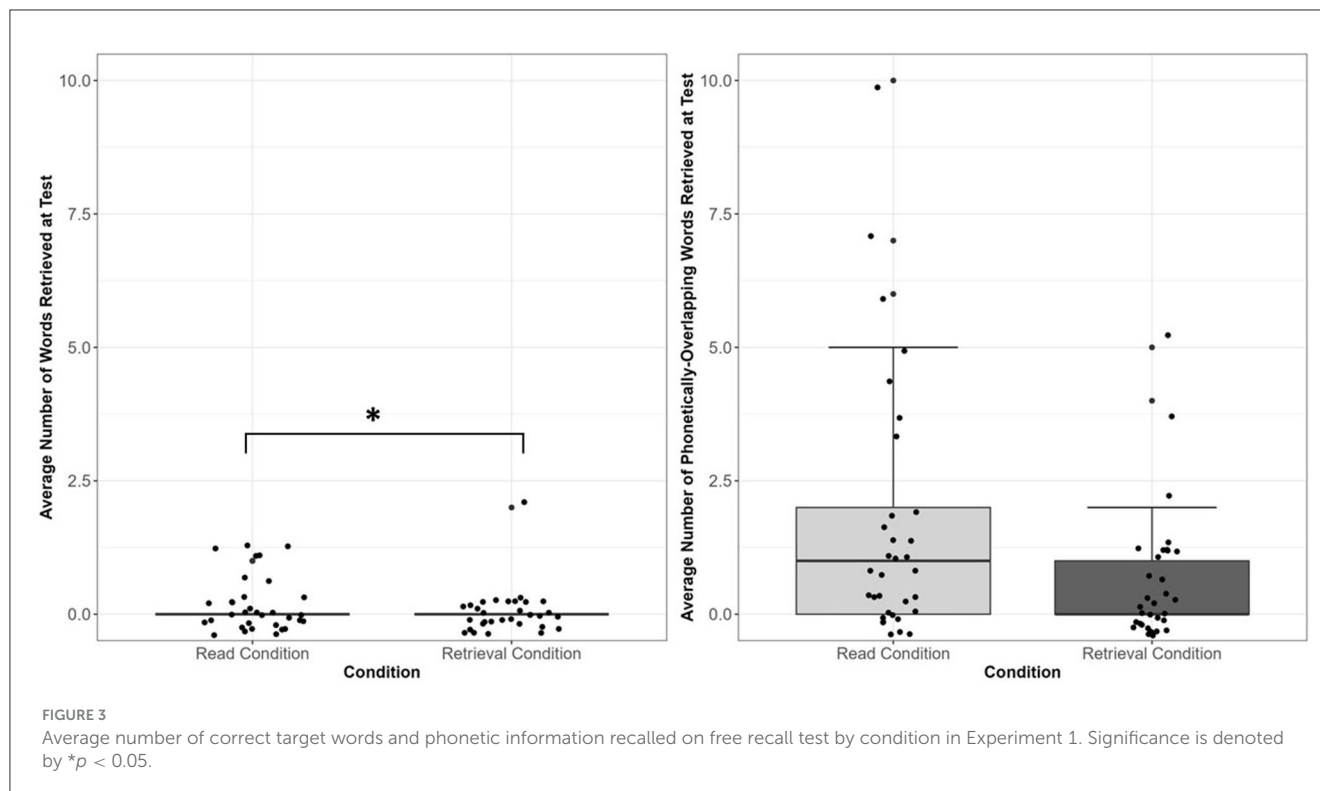
Transparency and openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures. All data and analyses for Experiments 1–3 are available at <https://osf.io/y6ges/>. This study's design and its analysis were not pre-registered.

Results and discussion

The aim of Experiment 1 was to assess whether a retrieval practice effect would emerge with a recall memory task. We predicted that retrieval practice would enhance memory for the novel words in comparison to restudying. We began our analyses by examining whether children in the Retrieval Condition successfully retrieved the novel words during the Learning Phase. After all, a central assumption of research on retrieval practice is that an optimal amount of difficulty is necessary to boost learning (Bjork, 1994; Bjork and Bjork, 2009; Roediger and Butler, 2011). Thus, if the effort necessary to retrieve words is too low or too high in this study, a retrieval practice effect may not be observed. The number of words retrieved out of 30 trials during the Learning Phase served as the outcome variable. Because word retrieval violated normality, $W = 0.86$, $p < 0.001$, a one-sample Wilcoxon signed rank test was used. This analysis revealed that recall performance during learning was significantly above 0 ($M = 21.06$, $SD = 9.13$), $V = 528$, $p < 0.001$. These results suggest that children were able to retrieve the new words during the Learning Phase.

Our next question was whether recall success during learning led to better post-test performance for children in the Retrieval Condition. To answer this question, we compared post-test performance between the experimental conditions. Performance was measured as the number of correct trials out of 10 total trials



on the recall test. The number of correct trials was right-skewed and violated normality as measured by a Shapiro–Wilk normality test in the Read Condition, $W = 0.50$, $p < 0.001$, and Retrieval Condition, $W = 0.17$, $p < 0.001$. A Mann–Whitney U -test was therefore used to assess differences in recall performance at post-test between the Read Condition ($M = 0.21$, $SD = 0.42$) and Retrieval Condition ($M = 0.06$, $SD = 0.35$; Figure 3). This analysis revealed significantly higher recall performance in the Read Condition in comparison to the Retrieval Condition, $W = 620$, $p = 0.035$.

These findings did not support our prediction: even though children who engaged in retrieval practice successfully recalled words during learning, they did not outperform children who engaged in restudying. Moreover, memory performance across both groups was low, with children recalling an average of 0–2 words out of 10. One explanation for this low performance is that learning was defined as recalling the entire word accurately. That is, if a child produced the word “wuk,” as opposed to “wug,” they would be scored as not knowing the word. Might the results differ if we considered the overlap between the target words and the onset phonemes of the words children produced?

To answer this question, we assessed the number of produced words that overlapped phonologically with the target words out of 10 trials. The number of these words was again right-skewed and violated a Shapiro–Wilk normality test in the Read Condition, $W = 0.71$, $p < 0.001$, and Retrieval Condition, $W = 0.58$, $p < 0.001$. However, a Mann–Whitney U -test revealed that the number of phonetically-similar words retrieved was not significantly different in the Read Condition ($M = 1.61$, $SD = 2.41$) vs. the Retrieval Condition ($M = 0.62$, $SD = 1.16$), $W = 653$, $p = 0.07$ (Figure 3). These results demonstrate that children in both conditions partially

TABLE 1 Simple linear regression analysis for Experiment 1.

Variable	b	SE_b	β
Intercept	0.009	0.036	
Age	−0.00	0.001	0.012
Proportion of retrieval successes	−0.007	0.022	−0.06
R^2	0.004		
Adjusted R^2	−0.065		
No. of obs.	31		

Outcome variable is proportion of correct responses on free recall post-test.

encoded the word forms. Although children who restudied the words had higher recall memory for the phonetic information due to the additional encoding opportunities (Larsen-Freeman, 2012), this difference did not reach significance.

Another question that emerged from these findings is whether retrieval success during learning predicted retrieval success at test. The results of a simple linear regression analysis demonstrated that retrieval success during learning did not significantly predict recall success at test (Table 1). This means that children who successfully retrieved the novel words throughout the storybook readings did not perform better on the final recall memory task than children who did not successfully retrieve the novel words during learning. One explanation for this finding is that post-test performance was not high enough to capture meaningful variation between children.

Taken together, our findings from Experiment 1 did not align with prior research that has observed better memory performance for children who engage in retrieval practice (Hudson, 1990;

Fritz et al., 2007; Ritchie et al., 2013). Even when we considered phonological overlap between the target words and produced words, children's recall performance at test was low. These results align with other studies that have shown low recall of word forms on production tasks (e.g., Leonard et al., 2021). Providing children with additional exposure to the words and feedback on their responses might be a promising strategy for increasing retrieval success (see Fazio and Marsh, 2019 for similar recommendations). Indeed, prior research has found that providing feedback can enhance the benefits of retrieval practice (Pashler et al., 2007; Roediger and Butler, 2011). The goal of Experiment 2 was therefore to include additional word mapping and retrieval opportunities during storybook reading.

Experiment 2

Experiment 2 examined whether a retrieval practice effect could be elicited if additional mapping and retrieval opportunities were provided during storybook reading. All children were read the storybook once. After the first reading, half of the children were given the words for the novel objects before being asked to recall the words again (Naming + Question). The other half of the children were asked to recall the words for the novel objects before being given the words again (Question + Naming). After a delay, children's memory was assessed using a recall memory task, in which they were asked to produce the word for a target object. Based on Experiment 1, we predicted that children who heard the words named before engaging in retrieval practice would outperform the children who engaged in retrieval practice first, as this would increase retrieval success during learning.

Method

Participants

The participants were 52 typically developing preschool-aged children ($M_{\text{age}} = 47.20$, $SD = 11.33$; range: 27–68 months, 27 girls) recruited from local day care centers and preschools. The participants came from a convenience sample or primarily middle- to upper-socioeconomic status households. Children received a storybook as a thank you for their participation.

Apparatus and stimuli

The experimental stimuli were similar to Experiment 1.

Design

The design was similar to Experiment 1. However, children in this experiment were presented with the storybook learning material in either a Question + Naming Condition ($n = 26$) or a Naming + Question Condition ($n = 26$).

Procedure

The procedure for Experiment 2 was similar to Experiment 1, with a notable difference in the experimental conditions (Figure 4).

Learning phase

Children in both conditions were introduced to the labels on the first reading via the experimenter pointing to and labeling the object (i.e., "I'll put the *wug* on the cake. This one is the *wug*."). On readings 2–4 in the Naming + Question Condition, the experimenter pointed to and labeled the object before asking the child to retrieve the label (i.e., "I'll put this one on the cake. This one is the *wug*. What is this one called?"). In the Question + Naming Condition, the experimenter asked the child to retrieve the label before pointing to and labeling the object (i.e., "I'll put this one on the cake. What is this one called? This one is the *wug*."). Children's responses to the retrieval test were recorded by the experimenter and the experimenter did not let the children know whether their retrieval attempt was correct or incorrect.

Retention phase

The retention phase was identical to Experiment 1.

Testing phase

The testing phase was identical to Experiment 1.

Results and discussion

The aim of Experiment 2 was to assess whether a retrieval practice effect would emerge in a recall memory task if additional mapping and retrieval opportunities were provided during learning. We predicted that children who heard the novel objects named before retrieval (Naming + Question Condition) would outperform children who retrieved the names first (Question + Naming Condition), as this would increase the likelihood of retrieval success. We began our analyses by examining whether children in both conditions successfully retrieved the novel words during the Learning Phase. Performance was calculated as the proportion of correct trials out of 30 trials. Because word retrieval was again skewed and violated a Shapiro–Wilk test of normality for the Naming + Question Condition, $W = 0.75$, $p < 0.001$, and Question + Naming Condition, $W = 0.78$, $p = 0.018$, a one-sample Wilcoxon signed rank test was used. This analysis revealed that recognition performance during learning was significantly above 0 for the Naming + Question Condition ($M = 27.88$, $SD = 2.98$), $V = 351$, $p < 0.001$, and the Question + Naming Condition ($M = 25.31$, $SD = 5.56$), $V = 351$, $p < 0.001$. These results demonstrate that additional retrieval and naming supports did boost children's retrieval success during learning.

Next, we examined whether post-test performance differed between the two conditions. Performance was calculated as the number of correct trials out of 10 total trials on the recall test. The number of correct trials was right-skewed and violated normality in the Naming + Question Condition, $W = 0.67$, $p < 0.001$, and the Question + Naming Condition, $W = 0.45$, $p < 0.001$. A Mann–Whitney U -test did not reveal any differences in recall performance at post-test between the Naming + Question Condition ($M = 0.46$, $SD = 0.71$) and Question + Naming Condition ($M = 0.19$, $SD = 0.49$), $W = 404$, $p = 0.110$ (Figure 5).

Much like Experiment 1, we assessed whether retrieval success during learning predicted retrieval success at test. The results of a simple linear regression analysis demonstrated that retrieval

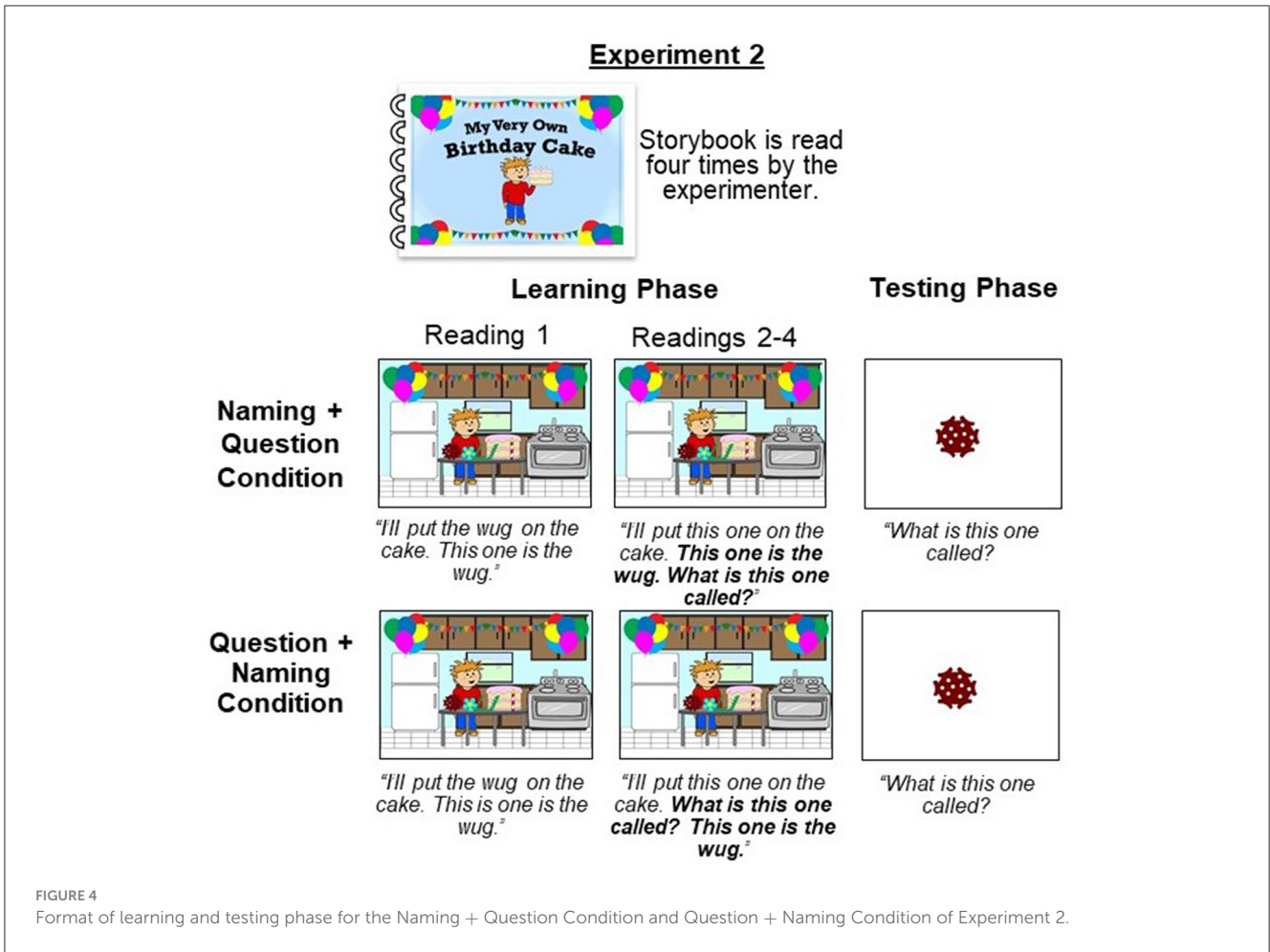


FIGURE 4 Format of learning and testing phase for the Naming + Question Condition and Question + Naming Condition of Experiment 2.

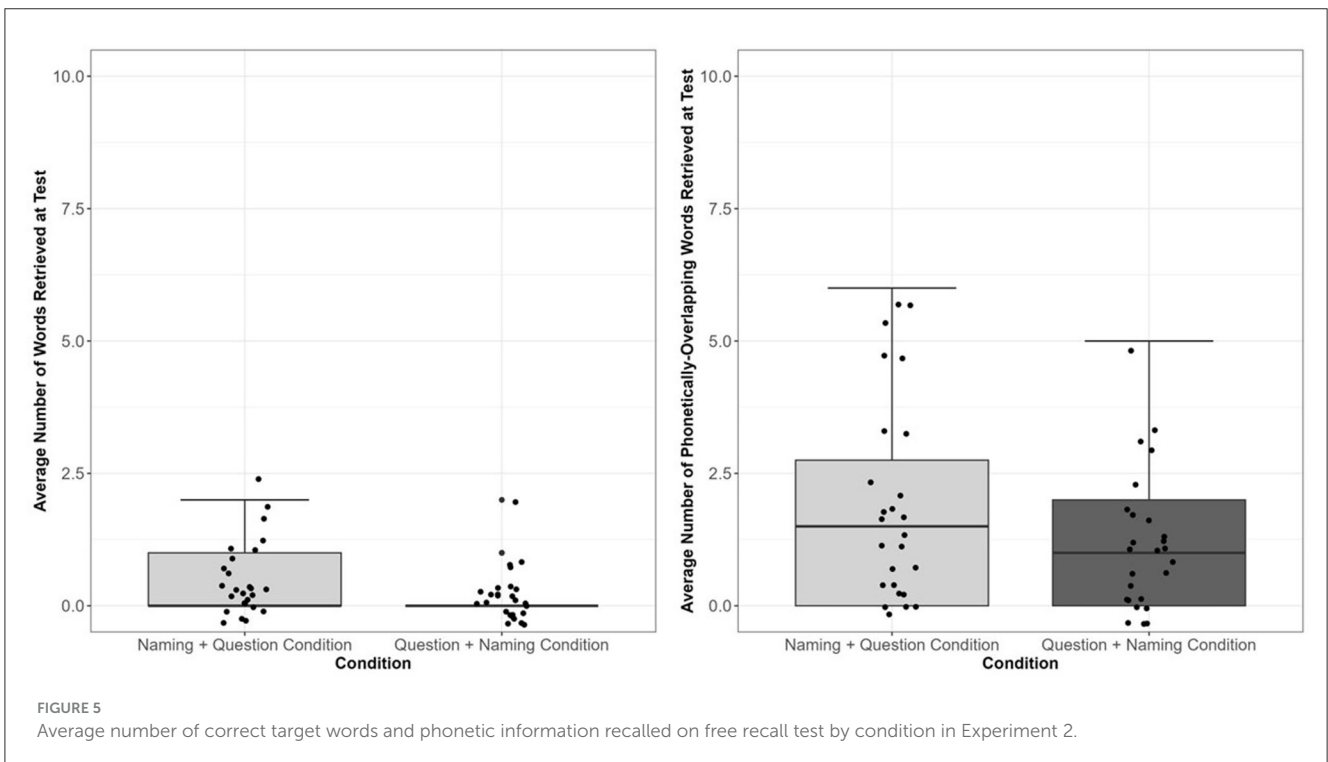


FIGURE 5 Average number of correct target words and phonetic information recalled on free recall test by condition in Experiment 2.

TABLE 2 Simple linear regression analysis for Experiment 2.

Variable	<i>b</i>	SE _{<i>b</i>}	β
Naming + Question			
Intercept	0.194	0.146	
Age	0.001	0.001	0.156
Proportion of retrieval successes	-0.212	0.152	-0.30
R ²	0.098		
Adjusted R ²	0.008		
No. of obs.	22		
Question + Naming			
Intercept	-0.047	0.074	
Age	0.00	0.001	0.067
Proportion of retrieval successes	0.063	0.067	0.204
R ²	0.047		
Adjusted R ²	-0.048		
No. of obs.	22		

Outcome variable is proportion of correct responses on forced-choice recognition post-test.

success during learning did not significantly predict retrieval success at test, controlling for age (Table 2). These results resemble the null findings from Experiment 1; in both experiments, the final post-test was a production task that yielded low performance and therefore low variation between children.

Because the post-test involved a recall memory task, we also assessed whether the produced words shared the same phoneme onset as the target words. Performance was calculated as the proportion of produced words that overlapped phonologically with the target words out of 10 trials. The number of these words was again right-skewed and violated a Shapiro-Wilk normality test in the Naming + Question Condition, $W = 0.84$, $p < 0.001$, and Question + Naming Condition, $W = 0.83$, $p < 0.001$. A Mann-Whitney U -test was therefore used to assess differences between groups and revealed no significant difference in the number of phonetically-similar words retrieved between the Naming + Question Condition ($M = 1.93$, $SD = 1.98$) vs. Question + Naming Condition ($M = 1.19$, $SD = 1.26$), $W = 399$, $p = 0.25$ (Figure 5).

The results of Experiment 2 again did not support our prediction: Children who heard the novel words before being asked to retrieve them did not outperform children who retrieved the words first. This was surprising because overall retrieval success during learning was very high for the Naming + Question Condition. Children in this condition retrieved on average 9.19 out of 10 target words ($SD = 1.04$) in the second read-through, 9.34 target words ($SD = 1.08$) in the third read-through, and 9.68 target words ($SD = 0.73$) in the final read-through of the storybook. Why did children in the Naming + Question Condition not outperform children in the Question + Naming Condition at test? Prior work has found that retrieval attempts—even when unsuccessful—can benefit learning if feedback is provided (e.g., Richland et al., 2009; Roediger and Butler, 2011; Grimaldi and Karpicke, 2012; Kornell, 2014; Kornell and Vaughn, 2016). That is, correctly remembering the word (a successful retrieval attempt) and receiving feedback

after an unsuccessful retrieval attempt can have the same benefit for learning. This might explain why performance did not differ significantly between the two conditions.

In summary, the results from Experiment 2 demonstrate that a retrieval practice effect did not emerge with additional mapping and retrieval opportunities, regardless of how these supports were provided and children's retrieval success during learning. These findings suggest that simply providing additional memory supports during learning is not sufficient to bolster children's recall at test. However, it is possible that the final recall task was too difficult to elicit a retrieval practice effect in Experiment 1 and 2. Prior studies have demonstrated that children struggle to recall words and other learned information after short time delays (e.g., Cole et al., 1971; Jablonski, 1974; Tversky and Teiffer, 1976; Dirks and Neisser, 1977; Cohen and Stewart, 1982; Ackerman, 1984; Horst and Samuelson, 2008; Vlach et al., 2008). In contrast, young children are better at recognizing previously learned information (e.g., Tversky and Teiffer, 1976; Brown et al., 1982). A retrieval practice effect might therefore be observed with a recognition memory task instead of a recall memory task. The aim of Experiment 3 was to test this possibility.

Experiment 3

Experiment 3 examined whether a retrieval practice effect could be elicited for children's recognition memory of novel words in the absence of additional mapping and retrieval opportunities during learning. Across two conditions, children engaged in passive word mapping or retrieval practice. Children's retrieval success was again measured during learning. After a delay, children's memory was assessed using a recognition memory task, in which they were asked to select a target object among an array of objects. We predicted that children who engaged in retrieval practice during storybook reading would outperform children who engaged in passive word mapping.

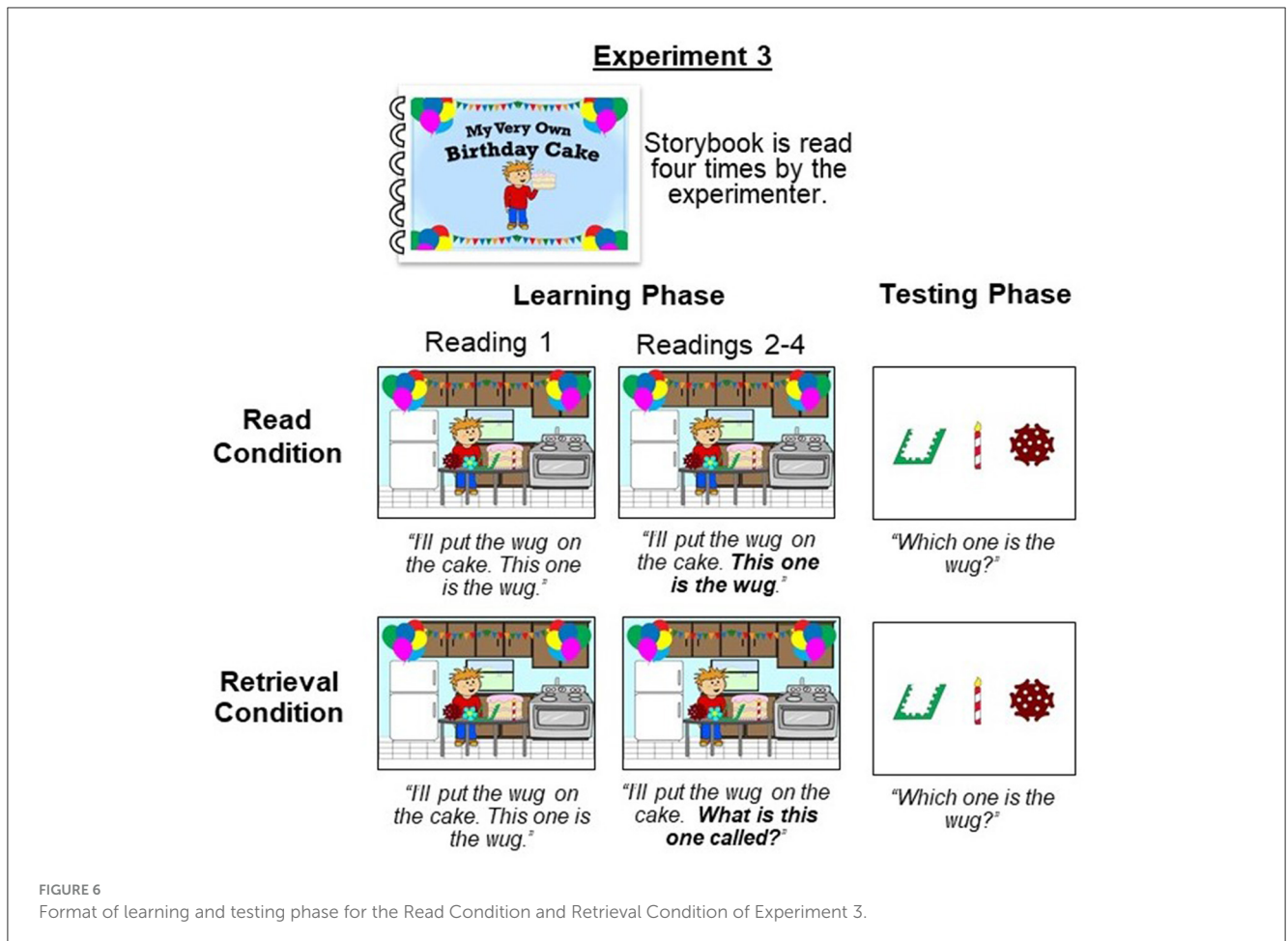
Method

Participants

The participants were 50 typically developing preschool-aged children ($M_{\text{age}} = 47.94$, $SD = 11.67$; range: 25–59 months, 19 girls) recruited from local daycare centers and preschools. An additional 14 children participated in the study but were not included in the final sample because of fussiness and inability to follow directions during the experiment (e.g., repeatedly choosing the object on the left side of the screen). The participants came from a convenience sample or primarily middle- to upper-socioeconomic status households. Children received a storybook as a thank you for their participation.

Apparatus and stimuli

The experimental stimuli were similar to Experiments 1–2, except for the testing phase trials: rather than producing a label after seeing an image of the target object (as was the procedure



in Experiment 1 and 2), children heard the label and selected the corresponding target object from among three objects.

Design

The design was similar to Experiment 1. Children in this experiment were presented with the storybook learning material in either a Passive Reading Condition ($n = 25$) or a Retrieval Practice Condition ($n = 25$).

Procedure

The procedure was similar to Experiment 1, with a notable difference in the testing trials (Figure 6).

Learning phase

The learning phase was identical to Experiment 1.

Retention phase

The retention phase was identical to Experiment 1.

Testing phase

Following the 5-min retention interval, the experimenter presented children with a forced-choice recognition test comprised of 10 trials. Each trial included a target novel object, a distractor novel object, and a familiar object from the story. The experimenter

asked children to identify the target novel object from the three objects presented on screen by saying, "Which one is the *wug*?". All three objects had appeared together in a scene and chance performance was at 33.33%. The experimenter did not provide feedback throughout testing. Children's answers were recorded on a response sheet. If no answer was selected, the trial was left blank and scored as incorrect.

Results and discussion

The aim of Experiment 3 was to assess whether a retrieval practice effect would emerge with a recognition memory task. We predicted that retrieval practice would enhance memory performance for the novel words in comparison to restudying. First, we examined whether children in the Retrieval Condition successfully retrieved the novel words during the Learning Phase; an important requirement for observing retrieval practice effects at test. Performance was calculated as the number of correct trials out of 30 trials during the Learning Phase. Because the number of words retrieved during the Learning Phase violated a Shapiro–Wilk test of normality, $W = 0.90$, $p = 0.018$, a one-sample Wilcoxon signed rank test was used to assess children's word retrieval. This analysis revealed that recognition performance during learning was significantly above 0 ($M = 22.44$, $SD = 6.53$), $V = 325$, $p <$

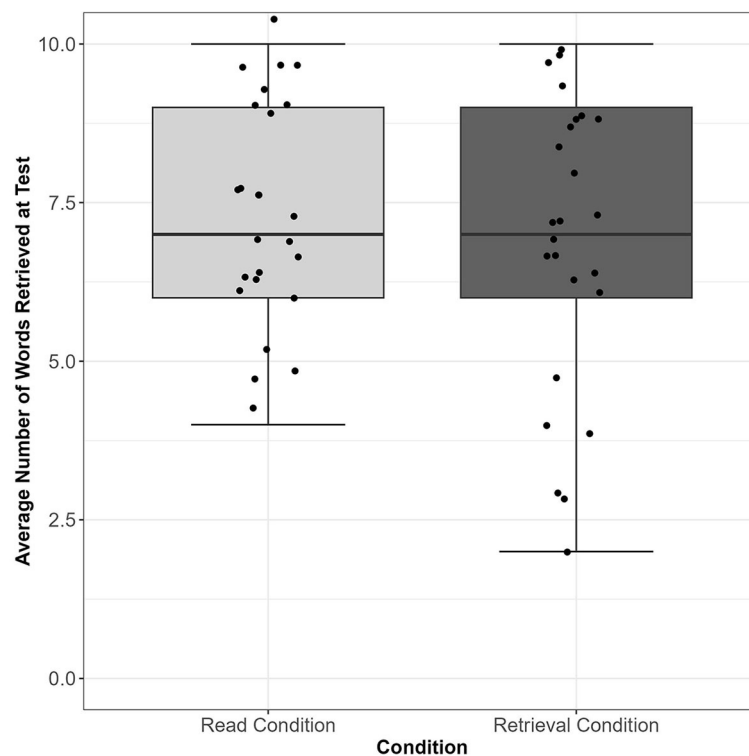


FIGURE 7
Average number of correct responses on forced-choice recognition post-test by condition in Experiment 3.

TABLE 3 Simple linear regression analysis for Experiment 3.

Variable	b	SE_b	β
Intercept	-0.021	0.158	
Age	0.002	0.003	0.102
Proportion of retrieval successes	0.803*	0.149	0.748
R^2	0.617		
Adjusted R^2	0.582		
No. of obs.	25		

Outcome variable is proportion of correct responses on forced-choice recognition post-test.
* $p < 0.05$.

0.001, suggesting that children were able to retrieve the new words during learning.

Next, we examined whether retrieval practice during learning led to better post-test performance for children in the Retrieval Condition. To answer this question, we compared post-test performance between the experimental conditions. Performance was measured as the number of correct trials out of 10 total trials on a recognition memory test. A Shapiro–Wilk normality test revealed that the normality assumption was not violated in the Read Condition, $W = 0.94$, $p = 0.13$, and Retrieval Condition, $W = 0.93$, $p = 0.08$. We therefore compared the mean number of words recognized between conditions using an independent samples t -test. This analysis revealed no significant difference in post-test performance between the Read Condition ($M = 7.40$, $SD = 1.80$)

and Retrieval Condition ($M = 6.88$, $SD = 2.33$), $t_{(48)} = 0.882$, $p = 0.382$, $d = 0.24$ (Figure 7).

Much like Experiments 1–2, our prediction was not supported: children who engaged in retrieval practice did not outperform children who engaged in restudying, even though they showed retrieval success during learning. Unlike Experiments 1–2, however, the results of a simple linear regression analysis demonstrated that retrieval success during learning significantly predicted retrieval success at test, controlling for age (Table 3). This means that children who could successfully retrieve the novel words throughout the storybook readings also performed better on the final recognition memory task. One explanation for this finding is that retrieval success at test was high enough to capture meaningful variation between children. Nonetheless, children did not demonstrate a retrieval practice effect when tested using a less demanding recognition memory task as opposed to a free recall test. The implications of these findings are outlined in the General Discussion.

General discussion

Increasing retrieval opportunities during learning leads to stronger long-term memory in older children and adults (Roediger and Karpicke, 2006; McDaniel et al., 2007; Marsh et al., 2012). The aim of the present study was to assess whether retrieval practice could also support preschoolers' memory for novel words in the context of storybook readings. Experiments 1–3 revealed

that retrieval practice did not provide an advantage over re-reading for children's recall and recognition memory of novel words at test, even though children could successfully map and retrieve the novel words during the storybook readings. Their success during learning was especially pronounced in Experiment 2 when additional retrieval supports (i.e., additional naming opportunities or feedback) were provided.

Despite high retrieval success during learning in all three experiments, we found null effects at test: Children struggled to retrieve the novel words, regardless of the test format (i.e., recall vs. recognition test). This study is not the first to find null effects in children. Studies with preschoolers and kindergarteners have found that retrieval practice does not benefit memory performance (Carneiro et al., 2018). Moreover, studies with school-aged children have found that young learners struggled to recall educational texts, even when the texts were available to them throughout testing (Karpicke et al., 2014). Performance was only improved after substantial memory supports were provided (i.e., children filled out a partially completed concept map). However, too much scaffolding at test can also lead to null results: in one study, retrieval practice benefited a fill-in-the-blank vocabulary test but did not benefit performance on a multiple-choice test due to ceiling effects (Goossens et al., 2014). Why does retrieval practice fail to improve children's learning in these situations?

To answer this question, we must consider why retrieval practice is effective for learning outcomes. Retrieval is an example of a cognitively effortful process that can be viewed as the bottleneck of the memory system. Whereas the capacity to store information in long-term memory is theoretically infinite, the capacity for retrieval is limited (Bjork and Bjork, 1992; Anderson and Neely, 1996). Retrieval is therefore effortful, yet learners increase the likelihood that they will retrieve information at a later point when they have difficulty accessing the information initially (Bjork, 1988; Bahrick and Hall, 2005; Pyc and Rawson, 2009). However, how difficult it is to access this information is related to how well it was learned, or stored, initially (Bjork and Bjork, 1992; McDaniel and Butler, 2011).

This distinction between storage and retrieval strength leads to two specific—thus far untested—hypotheses as to why retrieval practice might fail to enhance learning in young children. One hypothesis is that young children experience memory constraints, such as poor free recall and retrieval abilities, that hamper retrieval practice effects (Ghetti and Lee, 2011). We tested this conjecture in two ways: first, we incorporated additional retrieval supports in Experiment 2. Hearing the novel words named again and engaging in additional retrieval practice did increase retrieval success during learning but did not enhance learning at test. One explanation for this finding is that the frequency of retrieval attempts was increased, but not the spacing between learning and retrieval practice; that is, children engaged in retrieval practice immediately after hearing the novel words and not at a delay. Prior work has demonstrated that spacing out retrieval opportunities improves retention in comparison to providing immediate retrieval practice (Karpicke and Roediger, 2007; Haebig et al., 2019; Leonard and Deevy, 2020; Leonard et al., 2021, 2022). Spacing out retrieval practice in time might have engaged the retrieval processes necessary to successfully retrieve the words at test.

Another explanation for the poor recall at test is that children might have forgotten the words between the Learning and Testing Phases. After all, prior work has found that children rapidly forget verbal information across short periods of time (e.g., Brainerd et al., 1990; Ghetti and Angelini, 2008; Vlach and Sandhofer, 2012). Because the words in the storybook were novel, children might not have retained them long enough to retrieve them at test. To slow forgetting and enhance retrieval at test, future studies should improve how well the novel words are stored (their storage strength) by including fewer words or providing additional exposures to each word.

Second, we administered a forced-choice recognition memory test in Experiment 3 to see whether retrieval practice effects would emerge with a test format that posed fewer memory demands; however, this change to the test format did not lead to a retrieval practice effect.

Another hypothesis for why there was an absence of a retrieval practice effect is that young children differ from adults in their prior knowledge (e.g., Bjorklund, 1987; Shing and Brod, 2016; Knabe and Vlach, 2020); children are new learners and have acquired less knowledge than older children and adults. Studies have shown that the amount of prior knowledge affects how well newly learned information is stored and retrieved (Chi, 1978; Schneider et al., 1993; Elischberger, 2005; Brod et al., 2013). Thus, retrieval practice effects may only be observed when learners have rich semantic networks for the to-be-learned information. Verbal recall was likely challenging in the present study because children had to learn completely novel words (e.g., *wug*, *dax*, *fep*, *nola*, *blicket*, *gips*, *motch*, *boam*, *toma*, *hap*). This means they had no prior experience with these words other than their familiarity with English phonotactics. Thus, the results might have differed if a cued-recall technique had been used for the recall test (i.e., “What is this one called? It's a bli...”) or the study had included familiar, early-acquired words (e.g., *ball*, *dog*, *shoe*), with higher initial storage strength.

There are other possible explanations for the null results beyond learners' prior knowledge: retrieval practice effects may disappear for children in organic learning situations that involve complex stimuli (e.g., storybooks) and dynamic social interactions (e.g., parent-child dyads) (see Van Gog and Sweller, 2015, for a review on this issue in adults). Storybooks are complex learning tasks for preschoolers because they include a myriad of novel words, speakers, background scenes, an overarching narrative, and a reader. Indeed, research on novel word learning from storybooks has revealed that children learn associations beyond words and objects, and that word learning suffers when children track these additional associations (Knabe and Vlach, 2023). Together with their limited prior knowledge, complex learning situations might make retrieval practice too difficult for young children.

An important next step in research is therefore to assess when retrieval practice benefits young children's learning. For instance, researchers should use verbal stimuli that vary in difficulty (familiar words vs. novel words) to determine how prior knowledge affects retrieval practice. Retrieval practice should also be embedded in laboratory-based tasks (e.g., paired-associates task) and organic learning environments (e.g., preschool classrooms, parent-child dyads) to determine how the implementation of retrieval practice

influences memory performance. Finally, this work should be conducted in a broad age and language ability range to account for developmental differences in retrieval practice effects (see Haebig et al., 2019; Leonard and Deevy, 2020; Leonard et al., 2021, 2022 for work on retrieval practice in children with developmental language disorder). These endeavors will contribute to theoretical accounts that can explain developmental differences in retrieval practice effects. Indeed, existing theories explaining retrieval practice effects across different populations remain underdeveloped, even though retrieval practice is widely implemented in research and practice (Rowland, 2014).

In summary, researchers have claimed that retrieval practice is one of the most replicable learning phenomena in psychological science. For certain populations, such as college-aged students in university participant pools, and certain tasks, such as lab-based memory studies, that may be true. However, the current research found that preschoolers did not benefit from retrieval practice in a novel word learning paradigm, even after experiencing high retrieval success and receiving additional memory supports during learning. These results, combined with other null results in the literature, suggest that retrieval practice may only emerge with young children under specific conditions. Future research should therefore seek to identify the factors that predict when retrieval practice effects emerge across different populations. Critically, researchers should consider the developmental state of the learner and broader learning context when developing strategies to support memory retrieval.

Data availability statement

All data, analyses, and research materials are available at: <https://osf.io/y6ges/>. This study's design and its analysis were not pre-registered.

Ethics statement

The studies involving humans were approved by Social and Behavioral Sciences/Education Research IRB. The studies were conducted in accordance with the local legislation and institutional

requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

MK: Formal analysis, Visualization, Writing – original draft, Writing – review & editing. CB: Conceptualization, Data curation, Investigation, Methodology, Writing – review & editing. HV: Funding acquisition, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the James S. McDonnell Foundation Human Cognition Scholar Award to HV, as well as a Vilas Trust grant to HV.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ackerman, B. P. (1984). Item-specific and relational encoding effects in children's recall and recognition memory for words. *J. Exp. Child Psychol.* 37, 426–450. doi: 10.1016/0022-0965(84)90069-9
- Anderson, M. C., and Neely, J. H. (1996). "Interference and inhibition in memory retrieval," in *Memory*, eds E. L. Bjork and R. A. Bjork (Academic Press), 237–313. doi: 10.1016/B978-012102570-0/50010-0
- Bahrnick, H. P., and Hall, L. K. (2005). The importance of retrieval failures to long-term retention: a metacognitive explanation of the spacing effect. *J. Mem. Lang.* 52, 566–577. doi: 10.1016/j.jml.2005.01.012
- Bjork, E. L., and Bjork, R. A. (2009). "Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning," in *Psychology and the Real World*, eds M. A. Gernsbacher, R. W. Pew, and L. M. Hough (Gordonsville, VA: Worth), 56–64.
- Bjork, R. A. (1988). "Retrieval inhibition as an adaptive mechanism in human memory," in *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving*, eds H. L. Roediger, and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 309–330.
- Bjork, R. A. (1994). "Institutional impediments to effective training," in *Learning, Remembering, Believing: Enhancing Human Performance*, eds D. Druckman, and R. A. Bjork (Washington, DC: National Academy Press), 295–306.
- Bjork, R. A., and Bjork, E. L. (1992). "A new theory of disuse and an old theory of stimulus fluctuation," in *Essays in Honor of William K. Estes, Vol. 1. From Learning Theory to Connectionist Theory; Vol. 2. From Learning Processes to Cognitive Processes*, eds A. F. Healy, S. M. Kosslyn, and R. M. Shiffrin (Lawrence Erlbaum Associates, Inc.), 35–67.
- Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: an interpretive review. *Dev. Rev.* 7, 93–130. doi: 10.1016/0273-2297(87)90007-4

- Bleses, D., Makransky, G., Dale, P. S., Højen, A., and Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Appl. Psycholinguist.* 37, 1461–1476. doi: 10.1017/S0142716416000060
- Bouwmeester, S., and Verkoijen, P. P. J. L. (2011). Why do some children benefit more from testing than others? Gist trace processing to explain the testing effect. *J. Mem. Lang.* 65, 32–41. doi: 10.1016/j.jml.2011.02.005
- Brainerd, C. T., Reyna, V. F., Howe, M. L., Kingma, J., and Guttentag, R. E. (1990). The development of forgetting and reminiscence. *Monogr. Soc. Res. Child Dev.* 55, 1–109. doi: 10.2307/1166106
- Brod, G., Werkle-Bergner, M., and Shing, Y. L. (2013). The influence of prior knowledge on memory: a developmental cognitive neuroscience perspective. *Front. Behav. Neurosci.* 7, 1–13. doi: 10.3389/fnbeh.2013.00139
- Brown, A. L., Bransford, J. D., Ferrara, R. A., and Campione, J. C. (1982). “Learning, remembering, and understanding,” in *Carmichael’s Manual of Child Psychology*, eds J. H. Flavell, and E. M. Markman (New York, NY: Wiley Press), 80–84.
- Carneiro, P., Lapa, A., and Finn, B. (2018). The effect of unsuccessful retrieval on children’s subsequent learning. *J. Exp. Child Psychol.* 166, 400–420. doi: 10.1016/j.jecp.2017.09.010
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1563–1569. doi: 10.1037/a0017021
- Chi, M. T. H. (1978). “Knowledge structures and memory development,” in *Children’s Thinking: What develops?* 1st ed., ed. R. S. Siegler (London: Routledge), 75–96.
- Coane, J. H. (2013). Retrieval practice and elaborative encoding benefit memory in younger and older adults. *J. Appl. Res. Mem. Cogn.* 2, 95–100. doi: 10.1016/j.jarmac.2013.04.001
- Cohen, R. L., and Stewart, M. (1982). How to avoid developmental effects in free recall. *Scand. J. Psychol.* 23, 9–15. doi: 10.1111/j.1467-9450.1982.tb00408.x
- Cole, M., Frankel, F., and Sharp, D. (1971). Development of free recall learning in children. *Dev. Psychol.* 4, 109–123. doi: 10.1037/h0030435
- Dirks, J., and Neisser, U. (1977). Memory for objects in real scenes: the development of recognition and recall. *J. Exp. Child Psychol.* 23, 315–328. doi: 10.1016/0022-0965(77)90108-4
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students’ learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Public Interest* 14, 4–58. doi: 10.1177/1529100612453266
- Elishberger, H. B. (2005). The effects of prior knowledge on children’s memory and suggestibility. *J. Exp. Child Psychol.* 92, 247–275. doi: 10.1016/j.jecp.2005.05.002
- Endres, T., and Renkl, A. (2015). Mechanisms behind the testing effect: an empirical investigation of retrieval practice in meaningful learning. *Front. Psychol.* 6, 1054. doi: 10.3389/fpsyg.2015.01054
- Fazio, L. K., and Marsh, E. J. (2019). Retrieval-based learning in children. *Curr. Dir. Psychol. Sci.* 28, 111–116. doi: 10.1177/0963721418806673
- Flack, Z. M., Field, A. P., and Horst, J. S. (2018). The effects of shared storybook reading on word learning: a meta-analysis. *Dev. Psychol.* 54, 1334–1346. doi: 10.1037/dev0000512
- Fritz, C. O., Morris, P. E., Nolan, D., and Singleton, J. (2007). Expanding retrieval practice: an effective aid to preschool children’s learning. *Q. J. Exp. Psychol.* 60, 991–1004. doi: 10.1080/17470210600823595
- Ghetti, S., and Angelini, L. (2008). The development of recollection and familiarity in childhood and adolescence: evidence from the dual-process signal detection model. *Child Dev.* 79, 339–358. doi: 10.1111/j.1467-8624.2007.01129.x
- Ghetti, S., and Lee, J. (2011). Children’s episodic memory. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 365–373. doi: 10.1002/wcs.114
- Goossens, N. A. M. C., Camp, G., Verkoijen, P. P. J. L., Tabbers, H. K., and Zwaan, R. A. (2014). The benefit of retrieval practice over an elaborative restudy in primary school vocabulary learning. *J. Appl. Res. Mem. Cogn.* 3, 177–182. doi: 10.1037/h0101800
- Grimaldi, P. J., and Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Mem. Cogn.* 40, 505–513. doi: 10.3758/s13421-011-0174-0
- Haebig, E., Leonard, L. B., Deevy, P., Karpicke, J., Christ, S. L., Usler, E., et al. (2019). Retrieval-based word learning in young typically developing children and children with developmental language disorder II: a comparison of retrieval schedules. *J. Speech Lang. Hear. Res.* 62, 944–964. doi: 10.1044/2018_JSLHR-L-18-0071
- Horst, J. S., and Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month old infants. *Infancy* 13, 128–157. doi: 10.1080/15250000701795598
- Hudson, J. A. (1990). Constructive processing in children’s event memory. *Dev. Psychol.* 26, 180–187. doi: 10.1037/0012-1649.26.2.180
- Hudson, J. A., and Sheffield, E. G. (1998). Déjà vu all over again: effects of reenactment on toddlers’ event memory. *Child Dev.* 69, 51–67. doi: 10.1111/j.1467-8624.1998.tb06132.x
- Jablonski, E. M. (1974). Free recall in children. *Psychol. Bull.* 81, 522–539. doi: 10.1037/h0037012
- Jaeger, A., Eisenkraemer, R. E., and Stein, L. M. (2015). Test-enhanced learning in third-grade children. *Educ. Psychol.* 35, 513–521. doi: 10.1080/01443410.2014.963030
- Karpicke, J. D., and Aue, W. R. (2015). The testing effect is alive and well with complex materials. *Educ. Psychol. Rev.* 27, 317–326. doi: 10.1007/s10648-015-9309-3
- Karpicke, J. D., and Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 331, 772–775. doi: 10.1126/science.1199327
- Karpicke, J. D., Blunt, J. R., and Smith, M. A. (2016). Retrieval-based learning: positive effects of retrieval practice in elementary school children. *Front. Psychol.* 7, 350. doi: 10.3389/fpsyg.2016.00350
- Karpicke, J. D., Blunt, J. R., Smith, M. A., and Karpicke, S. S. (2014). Retrieval-based learning: the need for guided retrieval in elementary school children. *J. Appl. Res. Mem. Cogn.* 3, 198–206. doi: 10.1037/h0101802
- Karpicke, J. D., and Roediger, L. R. III (2007). Repeated retrieval during learning is the key to long-term retention. *J. Mem. Lang.* 57, 151–162. doi: 10.1016/j.jml.2006.09.004
- Knabe, M. L., and Vlach, H. A. (2020). When are difficulties desirable for children? First steps toward a developmental and individual differences account of the spacing effect. *J. Appl. Res. Mem. Cogn.* 9, 447–454. doi: 10.1016/j.jarmac.2020.07.007
- Knabe, M. L., and Vlach, H. A. (2023). Not all is forgotten: children’s memory for features of word learning episode. *Dev. Sci.* 26, e13291. doi: 10.1111/desc.13291
- Kornell, N. (2014). Attempting to answer a meaningful question enhances subsequent learning even when feedback is delayed. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 106–114. doi: 10.1037/a0033699
- Kornell, N., Hays, M. J., and Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *J. Exp. Psychol.* 35, 989–998. doi: 10.1037/a0015729
- Kornell, N., and Vaughn, K. E. (2016). How retrieval attempts affect learning: a review and synthesis. *Psychol. Learn. Motiv.* 65, 184–215. doi: 10.1016/bs.plm.2016.03.003
- Kucker, S. C., and Samuelson, L. K. (2012). The first slow step: differential effects of object and word-form familiarization on retention of fast-mapped words. *Infancy* 17, 295–323. doi: 10.1111/j.1532-7078.2011.00081.x
- Larsen-Freeman, D. (2012). On the roles of repetition in language teaching and learning. *Appl. Linguist. Rev.* 3, 195–210. doi: 10.1515/applrev-2012-0009
- Leonard, L. B., Christ, S. L., Deevy, P., Karpicke, J. D., Weber, C., Haebig, E., et al. (2021). A multi-study examination of the role of repeated spaced retrieval in the word learning of children with developmental language disorder. *J. Neurodev. Disord.* 13, 20. doi: 10.1186/s11689-021-09368-z
- Leonard, L. B., and Deevy, P. (2020). Retrieval practice and word learning in children with specific language impairment and their typically developing peers. *J. Speech Lang. Hear. Res.* 63, 3252–3262. doi: 10.1044/2020_JSLHR-20-00006
- Leonard, L. B., Kueser, J. B., Deevy, P., Haebig, E., Karpicke, J. D., Weber, C., et al. (2022). The contributions of immediate retrieval and spaced retrieval to word learning in preschoolers with developmental language disorder. *Autism Dev. Lang. Impair.* 7, 23969415221077652. doi: 10.1177/23969415221077652
- Maddox, G. B., and Balota, D. A. (2015). Retrieval practice and spacing effects in young and older adults: an examination of the benefits of desirable difficulty. *Mem. Cogn.* 43, 760–774. doi: 10.3758/s13421-014-0499-6
- Marsh, E. J., Fazio, L. K., and Goswick, A. E. (2012). Memorial consequences of testing school-aged children. *Memory* 20, 899–906. doi: 10.1080/09658211.2012.708757
- Masrai, A., and Milton, J. (2018). Measuring the contribution of academic and general vocabulary knowledge to learners’ academic achievement. *J. English Acad. Purp.* 31, 44–57. doi: 10.1016/j.jeap.2017.12.006
- McDaniel, M. A., Anderson, J. L., Debish, M. H., and Morrisette, N. (2007). Testing the testing effect in the classroom. *Eu. J. Cogn. Psychol.* 19, 494–513. doi: 10.1080/09541440701326154
- McDaniel, M. A., and Butler, A. C. (2011). “A contextual framework for understanding when difficulties are desirable,” in *Successful Remembering and Successful Forgetting: A Festschrift In Honor Of Robert A. Bjork*, ed. A. S. Benjamin (Psychology Press), 175–198.
- McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., and Roediger, H. L. (2013). Quizzing in middle-school science: successful transfer performance on classroom exams. *Appl. Cognit. Psychol.* 27, 360–372. doi: 10.1002/acp.2914
- Montag, J. L., Jones, M. N., and Smith, L. B. (2015). The words children hear: picture books and the statistics for language learning. *Psychol. Sci.* 26, 1489–1496. doi: 10.1177/0956797615594361
- Morgan, P. L., Farkas, G., Hillemeier, M. M., and Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educ. Res.* 45, 18–35. doi: 10.3102/0013189X16633182
- Ninio, A. (1983). Joint book reading as a multiple vocabulary acquisition device. *Dev. Psychol.* 19, 445–451. doi: 10.1037/0012-1649.19.3.445

- Ornstein, P. A., Haden, C. A., and Elischberger, H. B. (2006). "Children's memory development: remembering the past and preparing for the future," in *Lifespan Cognition: Mechanisms of Change*, eds E. Bialystok, and F. I. M. Craik (Oxford: Oxford University Press), 143–161. doi: 10.1093/acprof:oso/9780195169539.003.0010
- Pashler, H., Rohrer, D., Cepeda, N. J., and Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: choice and consequences. *Psychon. Bull. Rev.* 14, 187–193. doi: 10.3758/BF03194050
- Poole, D. A., and White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Dev. Psychol.* 27, 975–986. doi: 10.1037/0012-1649.27.6.975
- Pyc, M. A., and Rawson, K. A. (2009). Testing the retrieval effort hypothesis: does greater difficulty correctly recalling information lead to higher levels of memory. *J. Mem. Lang.* 60, 437–447. doi: 10.1016/j.jml.2009.01.004
- Richland, L. E., Kornell, N., and Kao, L. S. (2009). The pretesting effect: do unsuccessful retrieval attempts enhance learning? *J. Exp. Psychol. Appl.* 15, 243–257. doi: 10.1037/a0016496
- Ritchie, S. J., Sala, S. D., and McIntosh, R. D. (2013). Retrieval practice, with or without mind mapping, boosts fact learning in primary school children. *PLoS ONE* 8, e78976. doi: 10.1371/journal.pone.0078976
- Roediger III, H. L., Agarwal, P. K., McDaniel, M. A., and McDermott, K. B. (2011). Test-enhanced learning in the classroom: long-term improvement from quizzing. *J. Exp. Psychol. Appl.* 17, 382–395. doi: 10.1037/a0026252
- Roediger III, H. L., and Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends Cogn. Sci.* 15, 20–27. doi: 10.1016/j.tics.2010.09.003
- Roediger III, H. L., and Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249. doi: 10.1111/j.1467-9280.2006.01693.x
- Rohrer, D., Taylor, K., and Sholar, B. (2010). Test enhance the transfer of learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 233–239. doi: 10.1037/a0017678
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol. Bull.* 140, 1432–1463. doi: 10.1037/a0037559
- Schneider, W., Gruber, H., Gold, A., and Opwis, K. (1993). Chess expertise and memory for chess positions in children and adults. *J. Exp. Child Psychol.* 56, 328–349. doi: 10.1006/jecp.1993.1038
- Shing, Y. L., and Brod, G. (2016). Effects of prior knowledge on memory: implications for education. *Mind Brain Educ.* 10, 153–161. doi: 10.1111/mbe.12110
- Silverman, R. D. (2007). Vocabulary development of English-language and English-only learners in kindergarten. *Elem. Sch. J.* 107, 365–383. 0013-5984/2007/10704-0003\$05.00 doi: 10.1086/516669
- Tversky, B., and Teiffer, E. (1976). Development of strategies for recall and recognition. *Dev. Psychol.* 12, 406–410. doi: 10.1037/0012-1649.12.5.406
- Twomey, K. E., Ranson, S. L., and Horst, J. S. (2014). That's more like it: multiple exemplars facilitate word learning. *Infant Child Dev.* 23, 105–122. doi: 10.1002/icd.1824
- Van Gog, T., and Sweller, J. (2015). Not new, but nearly forgotten: the testing effect decreases or even disappears as the complexity of learning material increases. *Educ. Psychol. Rev.* 27, 247–264. doi: 10.1007/s10648-015-9310-x
- Vlach, H., and Sandhofer, C. M. (2012). Fast mapping across time: memory processes support children's retention of learned words. *Front. Psychol.* 3, 46. doi: 10.3389/fpsyg.2012.00046
- Vlach, H. A. (2019). Learning to remember words: memory constraints as double-edged sword mechanisms of language development. *Child Dev. Perspect.* 13, 159–165. doi: 10.1111/cdep.12337
- Vlach, H. A., Sandhofer, C. M., and Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition* 109, 163–167. doi: 10.1016/j.cognition.2008.07.013
- Wasik, B. A., Hindman, A. H., and Snell, E. K. (2016). Book reading and vocabulary development: a systematic review. *Early Child. Res. Q.* 37, 39–57. doi: 10.1016/j.ecresq.2016.04.003