Psychonomic Bulletin & Review (2026) 33:230
https://doi.org/10.3758/s13423-026-03004-4

BRIEF REPORT

# Selective retrieval boosts memory regardless of prior knowledge

**Zhehao Huang[1] · Yang Cao[1,2] · Jiali Li[1] · Geoffrey F. Woodman[3,4] · Bao Zhang[5] · Benchi Wang[2]**

Received: 17 November 2025 / Accepted: 18 August 2026
© The Psychonomic Society, Inc. 2026

**Abstract**
Forgetting is a core dynamic of memory which follows a rapid loss that slows into gradual decline. Notably, recent evidence suggests that selective retrieval strengthens recalled items, benefits nonretrieved ones, and reinstates the forgetting curve. A central question is whether these retrieval benefits depend on prior knowledge or reflect a domain-general mechanism. We tested 216 participants who studied high-knowledge (common) or low-knowledge (rare) Chinese characters and were assessed either directly or after a selective retrieval phase. Across conditions, retrieval boosted recall for both retrieved and nonretrieved characters and restarted forgetting functions. High-knowledge characters were remembered more accurately overall, but forgetting rates and retrieval-induced gains were equivalent across knowledge levels. These findings suggest that retrieval benefits can emerge independent of semantic expertise. By bridging expertise and forgetting literatures, our study provides evidence that selective retrieval as a domain-general phenomenon for stabilizing memory over time.

**Keywords** Memory · Prior knowledge · Retrieval · Forgetting

## Introduction

Memory performance typically declines as the interval between encoding and retrieval increases, with rapid initial loss followed by a more gradual decline over time (Ebbinghaus, 1885). Importantly, retrieving a subset of studied items after a delay may counteract this decline, enhancing memory not only for the retrieved items but, under some conditions, also for related nonretrieved items (Bäuml & Schlichting, 2014; Bäuml & Trißl, 2022). These effects can be examined using the selective retrieval paradigm, in which participants study a set of items and subsequently retrieve only a subset of them. This procedure yields two item classes: retrieved items that receive direct retrieval practice and nonretrieved items that remain untested but are associated with the same study episode. Comparing these item types makes it possible to assess how selective retrieval influences the retention of both directly accessed and related information over time. A central theoretical question is whether these retrieval benefits depend on prior knowledge or instead reflect a domain-general process that operates independently of semantic expertise.

This question sits at the intersection of two major literatures that make conflicting predictions. Research on expertise shows that schemas and semantic scaffolds shape memory, enabling experts to encode and retrieve more effectively

✉ Yang Cao
gnaycao.cy@gmail.com

✉ Bao Zhang
zhangbao@gzhu.edu.cn

✉ Benchi Wang
wangbenchi.swift@gmail.com

[1] Institute for Brain Research and Rehabilitation, South China Normal University, Guangzhou 510631, China

[2] Key Laboratory of Mental Health of the Ministry of Education, Guangdong-Hong Kong-Macao Greater Bay Area Center for Brain Science and Brain-Inspired Intelligence, Guangdong-Hong Kong Joint Laboratory for Psychiatric Disorders, Guangdong Province Key Laboratory of Psychiatric Disorders, Guangdong Basic Research Center of Excellence for Integrated Traditional and Western Medicine for Qingzhi Diseases, Department of Neurobiology, School of Basic Medical Sciences, Southern Medical University, Guangzhou, China

[3] Department of Psychology, Vanderbilt University, Nashville, TN, USA

[4] Center for Integrative and Cognitive Neuroscience, Vanderbilt University, Nashville, TN, USA

[5] School of Education, Guangzhou University, Guangzhou 510006, China

Published online: 08 September 2026

than novices (Sala & Gobet, 2017; Wing et al., 2022). From this view, retrieval benefits may manifest in two separable ways: an immediate improvement in recall performance and a potential alteration of the subsequent forgetting trajectory. Critically, these components are not necessarily coupled and must be assessed independently (Bäuml & Trißl, 2022; Roediger & Karpicke, 2006). Conversely, recent research on forgetting suggests that the rate of decay remains remarkably stable across manipulations of initial strength, repetition, and age (Rivera-Lares et al., 2022, 2023). Temporal-context accounts propose that memories are encoded together with features of the study context, which gradually drift over time and reduce accessibility. Selective retrieval may partially reinstate this context, thereby facilitating access to both retrieved and related nonretrieved items (Bäuml & Schlichting, 2014; Bäuml & Trißl, 2022; Polyn et al., 2009; Polyn & Woodman, 2025). To the extent that such retrieval benefits rely on contextual reinstatement rather than semantic scaffolding, they may emerge independently of prior knowledge.

We tested these alternatives using Chinese characters that naturally vary in familiarity. Common characters are processed with fluent semantic and phonological support, whereas rare characters require effortful structural analysis with minimal prior knowledge (Chen & Peng, 1998; Perfetti & Liu, 2005). Adapting the selective retrieval paradigm (Bäuml & Trißl, 2022), we compared the absolute forgetting curves for high-knowledge and low-knowledge characters, both with and without a retrieval phase. This design closely follows their experimental framework while extending it to examine how prior knowledge influences retrieval-induced changes in memory over time. Our primary analyses focus on absolute forgetting, defined as the reduction in recall performance over time and expressed as the number of items or percentage points lost. Equal absolute forgetting therefore indicates an equivalent amount of memory loss across conditions. By contrast, relative forgetting expresses this loss as a proportion of the initial recall level. Consequently, equal absolute forgetting does not necessarily imply equal relative forgetting[1].

This design isolates whether retrieval-induced enhancement is knowledge-dependent, as expertise accounts predict, or knowledge-independent, as temporal-context accounts propose. Resolving this issue clarifies not only the mechanisms of retrieval-based learning but also the architecture of forgetting itself, addressing whether retrieval practice serves as a phenomenon of memory stabilization or one constrained by prior knowledge.

[1] For example, declines from 80% to 60% and from 60% to 40% both represent an absolute loss of 20 percentage points but relative losses of 25% and 33.3%, respectively.

**Table 1** Stimuli used in the experiment

| | Knowledge | Chinese characters | Radicals |
|---|---|---|---|
| Practice | N/A | 全 鸡 网 妹 声 间 | 人 又 冂 女 士 门 |
| Formal test | High | 冰 但 夜 后 医 共<br>在 国 学 往 多 饭<br>有 现 或 所 武 妥 | 氵 亻 亠 厂 匚 八<br>土 口 子 彳 夕 饣<br>月 王 戈 斤 止 罒 |
| | Low | 陋 卣 凼 劬 兕 冏<br>弢 茉 庋 渤 邕 尪<br>炊 恣 庠 栻 昆 贬 | 阝 卜 凵 力 儿 冂<br>弓 艹 广 氵 山 尢<br>火 文 户 木 曰 见 |

Forty-two Chinese characters were selected based on familiarity ratings from an independent pretest. Eighteen high-knowledge (familiar) and 18 low-knowledge (rare) characters were used in the formal experiment, and six additional characters served as practice stimuli. Each character contained a unique radical used as a retrieval cue.

## Methods

### Participants

Two hundred and sixteen right-handed volunteers (148 women, aged 19.9 ± 1.4 years) were recruited from South China Normal University. Participants were randomly assigned to two groups (selective retrieval vs. nonretrieval), each containing four subgroups ($n = 27$) corresponding to different test delays (2, 10, 20, or 30 min). The sample size approximated prior studies investigating time-dependent forgetting and selective retrieval (subgroup $n = 24$–$28$; Bäuml & Dobler, 2015; Bäuml & Schlichting, 2014; Kliegl et al., 2019) and provided 0.82 power to replicate medium-sized effects. All participants had normal or corrected-to-normal vision, no neurological or cognitive disorders, and provided written informed consent.

### Materials and apparatus

Stimuli were derived from the Modern Chinese Dictionary (Dictionary Editing Office, Institute of Linguistics, Chinese Academy of Social Sciences, 2016). Characters were first screened to include radicals with two–four strokes and total stroke counts between six and eight, producing 1,547 candidates. To control phonological, orthographic and semantic factors, only characters with unique pronunciations, clear left-right, top-bottom, or enclosure structures, and differed semantically from its radical, were retained (Ding et al., 2000; Ferrand & Grainger, 2003; Taft et al., 2000; Yeh, & Li, 2002; Zhang et al., 2020).

Fifty independent raters (28 women; 20.5 ± 1.2 years) evaluated the familiarity of each character on a 7-point scale (1 = *very unfamiliar*, 7 = *very familiar*). Based on these ratings, 42 characters were selected: 36 for the main experiment and six for practice (see Table 1). The 36 experimental characters formed two lists—high-knowledge ($M$ = 6.84 ± 0.83, indicating high familiarity) and low-knowledge ($M$ = 1.67 0.63, indicating low familiarity)—each comprising 18 characters. Within each list, 10 characters served as *retrieved* items and 8 as *nonretrieved* items when selective retrieval was included, whereas all characters served as *studied* items when selective retrieval was absent (see Fig. 1). Stimuli were identical for all participants, and each character contained a unique radical serving as its retrieval cue.

Stimulus presentation was controlled by custom Python scripts. Participants placed their chins on a chin rest 60 cm from the monitor used to present the stimuli.

## Procedure and design

The design was a 2 (knowledge: high vs. low, within-subjects) × 2 (selective retrieval: absent vs. present, between-subjects) × 4 (delay: 2, 10, 20, or 30 min, between-subjects) mixed design. Figure 1 depicts the procedure for high-knowledge characters (the same applied to low knowledge).

During study, 18 characters were presented individually for 6 s each in random order. To prevent rehearsal and maintain comparable cognitive engagement across delays, delay intervals were filled with one or more neutral distracting tasks. These included a mental arithmetic task, a spatial cueing task, a visual search task, and a change-detection task, adapted from established paradigms (Posner & Cohen, 1984; Wang & Theeuwes, 2018; Wheeler & Treisman, 2002). For delay intervals of 20 min and 30 min, as well as for the 30-min interval preceding selective retrieval, time were subdivided into 10-min segments (two or three segments, respectively). Each segment was assigned a different distractor task, randomly selected without replacement from the spatial cueing, visual search, and change detection tasks, ensuring that no task was repeated within a participant's session. The mental arithmetic task was used only for the 2-min delay condition.

Selective retrieval began 30 min after study and involved two rounds of retrieval practice. In each round, participants attempted to recall 10 of the 18 characters (i.e., retrieved items) using their radicals as cues (6 s each) and wrote down the corresponding complete characters. Participants, who did not take part in selective retrieval, performed an equivalent neutral task instead.

At test, radicals of all 18 studied items were presented in random order. In the selective-retrieval group, nonretrieved items were always tested first to avoid potential reactivation spillover from retrieval of related items (Anderson et al., 1994; Bäuml & Dobler, 2015; Bäuml & Schlichting, 2014).

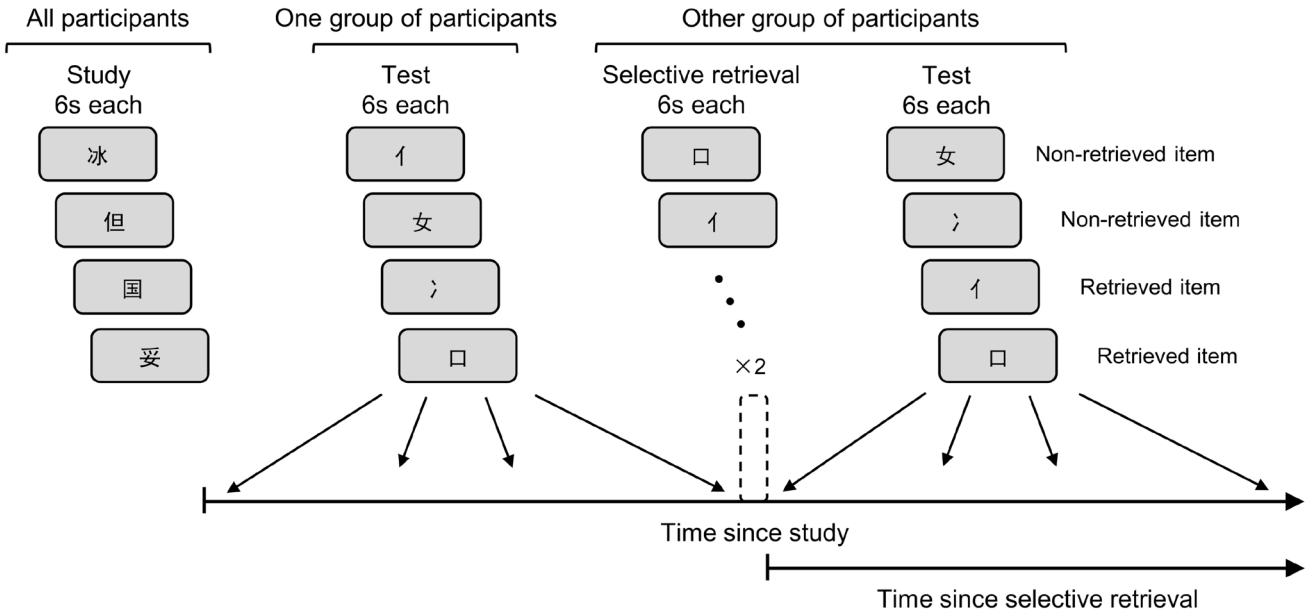



**Fig. 1** Experimental procedure. Participants studied 18 Chinese characters that varied in familiarity (high vs. low knowledge). In the control group, memory was tested after a delay of 2, 10, 20, or 30 min without any intermediate retrieval. In the selective retrieval group, participants completed two rounds of retrieval practice 30 min after study, producing retrieved and nonretrieved items. They were then tested after an additional 2, 10, 20, or 30 min. In the selective retrieval condition, nonretrieved items were always tested before retrieved items to avoid reactivation spillover

### Power-function fitting

To characterize the temporal dynamics of forgetting, we fitted a power function of time, $R(t) = at^{-b}$, to recall accuracy for each item types in different knowledge levels using maximum likelihood estimation (Bäuml & Triβl, 2022; Riefer & Batchelder, 1988). In this function, $R(t)$ denotes the percentage of items recalled at time $t$, parameter $b$ reflects the forgetting rate over time, and parameter $a$ represents the initial recall level. Consistent with prior work, functions were fitted to the group-average data (Wixted & Ebbesen, 1997).

## Results

### Forgetting without selective retrieval

We first examined how recall accuracy changed over time when no selective retrieval occurred. To verify that these baseline results were not confounded by output interference or testing order, we compared the recall accuracy of studied items tested in the first half of the sequence (Positions 1–8) with those tested in the second half (Positions 9–16). A 2 (testing position: first half vs. second half) × 4 (delay time: 2, 10, 20, or 30 min) repeated-measures ANOVA revealed no significant main effect of testing position for either high-knowledge, $F(1, 104) = 1.54$, $p = .22$, $\eta_p^2 = .02$, or low-knowledge items, $F(1, 104) = 3.27$, $p = .07$, $\eta_p^2 = .03$, and no significant interaction between testing position and delay time, high-knowledge: $F(3, 104) = .03$, $p = .99$, $\eta_p^2 = .001$; low-knowledge: $F(3, 104) = .41$, $p = .74$, $\eta_p^2 = .01$. These results indicate that recall performance remained stable across the test sequence. Accordingly, data were combined across testing positions. Overall, recall of these studied items declined systematically for both high-knowledge and low-knowledge characters (Fig. 2 and Table 2). A 2 (knowledge: high vs. low) × 4 (delay: 2, 10, 20, 30 min) repeated-measures ANOVA showed significant main effects of knowledge,

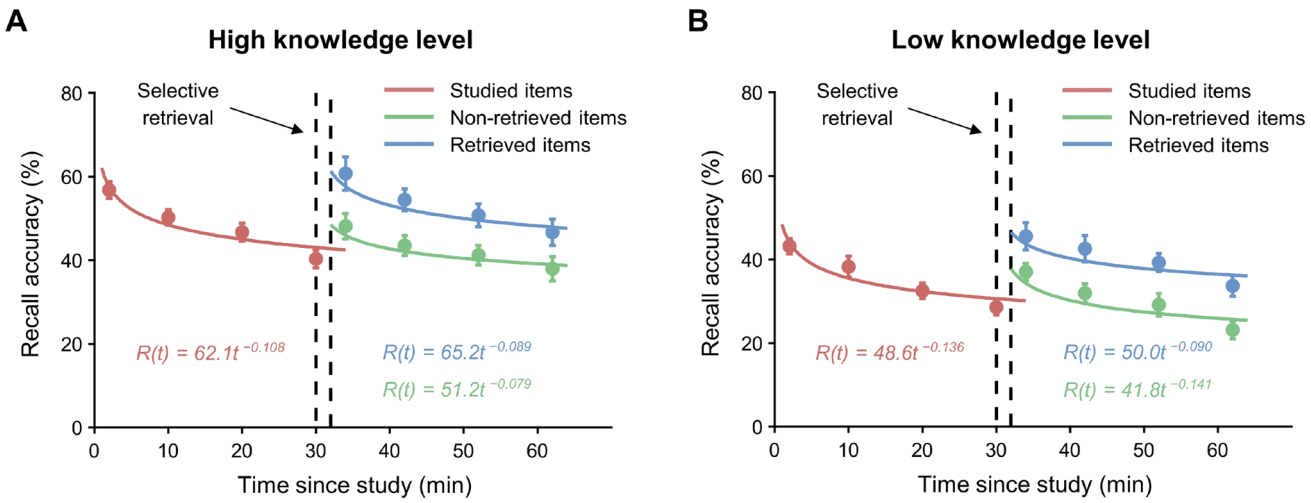



**Fig. 2** Recall accuracy of high-knowledge (**A**) and low-knowledge (**B**) characters across delay intervals. Dots indicates mean recall accuracy at the corresponding delay, and error bars show standard error of the mean (*SEM*). Lines show power-function fits, $R(t) = at^{-b}$, included for descriptive purposes to illustrate the temporal trajectory of forgetting across conditions. (Colour figure online)

**Table 2** Mean recall accuracy (%, $M \pm SEM$) across delay intervals

| Item | Knowledge | Delay (min) | | | |
|---|---|---|---|---|---|
| | | 2 | 10 | 20 | 30 |
| Studied | High | 56.79 ± 1.97 | 50.21 ± 1.79 | 46.71 ± 2.06 | 40.33 ± 2.09 |
| | Low | 43.21 ± 1.79 | 38.27 ± 2.42 | 32.50 ± 1.82 | 28.59 ± 1.89 |
| Retrieved | High | 60.74 ± 3.92 | 54.44 ± 2.58 | 50.74 ± 2.67 | 46.67 ± 3.11 |
| | Low | 45.56 ± 3.22 | 44.07 ± 3.47 | 39.26 ± 2.13 | 33.70 ± 2.45 |
| Nonretrieved | High | 48.15 ± 2.96 | 43.52 ± 2.35 | 41.20 ± 2.29 | 37.96 ± 2.87 |
| | Low | 37.04 ± 2.06 | 31.94 ± 2.25 | 29.17 ± 2.67 | 23.25 ± 2.64 |

Data are reported separately for studied, retrieved, and nonretrieved items at both knowledge levels (high vs. low).

$F(1, 104) = 103.5$, $p < .001$, $\eta_p^2 = 0.5$, and delay, $F(3, 104) = 18.6$, $p < .001$, $\eta_p^2 = 0.35$, with no interaction, $F(3, 104) = 0.20$, $p = .87$, $\eta_p^2 < .01$. Thus, forgetting followed the same temporal function for both knowledge levels: high-knowledge characters were recalled more accurately overall, but the rate of decline was equivalent. Consistently, planned comparisons showed that for high-knowledge items, recall after 30 min was significantly lower than after 2 min, $t(52) = 5.73$, $p < .001$; the same pattern appeared for low-knowledge items, $t(52) = 5.61$, $p < .001$.

Power-function fits provided a descriptive characterization of the temporal patterns of forgetting (Wixted & Ebbesen, 1997). Goodness-of-fit values indicated that the model captured the data well for studied, retrieved, and nonretrieved items across both knowledge levels (all $R^2$ values >.75). These fits illustrate that, although recall levels differed across conditions, the overall trajectory of decline over time was similar.

## Forgetting after selective retrieval

Next, we examined how selective retrieval modified this temporal forgetting. Retrieval-related improvements were assessed by comparing recall of retrieved and nonretrieved items at the 2-min delay following selective retrieval with recall of studied items at the 30-min delay in the no-retrieval condition. When selective retrieval occurred 30 min after study, recall improved sharply for both retrieved and nonretrieved items, indicating that retrieval interrupted ongoing forgetting (Fig. 2 and Table 2). Compared with control performance of studied items at an equivalent delay (i.e., 3 min), recall was significantly higher for retrieved items, $t(52) = 4.91$, $p < .001$ (high-knowledge) and $t(52) = 4.54$, $p < .001$ (low-knowledge). Nonretrieved items also showed reliable improvement, $t(52) = 2.20$, $p = .03$ (high-knowledge) and $t(52) = 3.02$, $p = .004$ (low-knowledge), though their accuracy remained lower than retrieved items, $t(26) = 2.78$, $p = .01$ (high-knowledge) and $t(26) = 2.28$, $p = .03$ (low-knowledge). Notably, these beneficial effects may be underestimated, given that the effective retention interval for retrieved and nonretrieved items was 34 min (30-min delay plus retrieval phase and 2-min delay), compared to the strict 30-min interval for the studied baseline. Importantly, these retrieval-induced gains did not differ significantly between high- and low-knowledge for retrieved items, $t(26) = 0.59$, $p = .56$ and nonretrieved items, $t(26) = 0.14$, $p = .89$. Thus, while selective retrieval enhanced memory performance overall, the magnitude of this benefit was comparable across knowledge levels.

A complementary comparison for recall performance at the 2-min delay showed that nonretrieved items were recalled less accurately than both studied and retrieved items, whereas retrieved and studied items did not differ significantly. In addition, high-knowledge characters showed significantly better performance than low-knowledge characters across item types, but no interaction between item type and knowledge level was observed, indicating that the pattern of retrieval-related differences was consistent across knowledge conditions (see Fig. S1 and supplement for details).

Despite these benefits, both retrieved and nonretrieved items showed clear time-dependent forgetting (Fig. 2). To verify that the retrieval-related differences reflected changes in overall recall level rather than differences in absolute forgetting rate, we directly compared forgetting across item types. A 3 (item type: studied, retrieved, nonretrieved) × 4 (delay: 2, 10, 20, 30 min) ANOVA was separately conducted for each knowledge level. For both high-knowledge and low-knowledge conditions, significant main effects of item type and delay were observed, but critically, no interaction between item type and delay was found ($p$ values >.95; see supplementary for details). This absence of interaction indicates that studied, retrieved, and nonretrieved items exhibited comparable rates of forgetting over time, despite differences in overall recall performance.

We then tested whether prior knowledge modulated temporal dynamics of forgetting after selective retrieval. Separate 2 (knowledge: high vs. low) × 4 (delay: 2, 10, 20, 30 min) repeated-measures ANOVAs provided evidence. For retrieved items, significant main effects were found for knowledge, $F(1, 104) = 47.01$, $p < .001$, $\eta_p^2 = 0.31$, and delay, $F(3, 104) = 5.85$, $p = .001$, $\eta_p^2 = 0.14$, with no interaction, $F(3, 104) = 0.2$, $p = .9$, $\eta_p^2 < .01$. For nonretrieved items, we observed main effects of knowledge, $F(1, 104) = 47.61$, $p < .001$, $\eta_p^2 = 0.31$, and delay, $F(3, 104) = 7.96$, $p < .001$, $\eta_p^2 = 0.19$, and again no interaction, $F(3, 104) = 0.22$, $p = .89$, $\eta_p^2 < .01$. Hence, after selective retrieval, forgetting proceeded at a similar rate for both knowledge levels, with high-knowledge items maintaining an overall advantage in recall accuracy. Consistently, planned comparisons further showed that recall after 30 min was significantly lower than after 2 min for both knowledge levels (retrieved: $t$ values > 2.81, $p$ values <.01; nonretrieved: $t$ values > 2.47, $p$ values <.05), confirming that selective retrieval modifies the forgetting curve rather than eliminating it.

## Relative forgetting

Our primary analyses focused on absolute forgetting. We next examined whether the same general pattern emerged when forgetting was expressed relative to recall performance. First, we compared the parameter $b$ derived from the power-function fits, which serves as a standard index of relative decay (Nickl & Bäuml, 2024). The estimated $b$ values were numerically similar across knowledge levels for each

item type (see Table S1), providing no descriptive evidence that prior knowledge systematically altered the proportional rate of forgetting.

We additionally calculated the relative loss between consecutive delay intervals for each knowledge level using the formula $(n_t - n_{t+1})/n_t$ ($n_t$ denotes the average number of items that participants recalled at delay $t$; Wixted, 2022). Due to our between-subjects design for the time delay, we used a bootstrap procedure with 10,000 iterations to estimate the 95% confidence interval (CI) of the difference in relative forgetting between high-knowledge and low-knowledge groups for the studied, retrieved and nonretrieved item types. The results showed that the 95% CI of the knowledge difference under three item types contained zero, indicating that the relative forgetting rates were statistically indistinguishable across knowledge levels (see Table S2 and Fig. S2). Together, the power-function estimates and interval-based analyses indicate that the general pattern observed for absolute forgetting was also evident when forgetting was expressed in relative terms.

## Discussion

This study tested whether the benefits of selective retrieval depend on prior knowledge. Across more than 200 participants, selective retrieval enhanced memory for both retrieved and nonretrieved characters and modified the forgetting curve into a new trajectory. Our primary analyses focused on absolute forgetting, defined as the amount of recall lost over time. Crucially, although high-knowledge characters yielded higher baseline recall, the rate of absolute forgetting and the magnitude of retrieval-induced gains were statistically equivalent across knowledge levels. Complementary analyses of relative forgetting yielded the same general pattern. These results provide direct evidence that retrieval benefits can emerge independently of prior knowledge, supporting a domain-general account rather than one reliant on semantic scaffolds.

The present findings refine our understanding of the role of preexisting semantic networks in memory enhancement. While prior knowledge improves overall recall performance, it does not appear to alter the dynamics of retrieval-induced changes in forgetting within the present paradigm. Extensive work shows that schemas and conceptual knowledge improve encoding and retrieval efficiency (Sala & Gobet, 2017; Wing et al., 2022). If selective retrieval relied on such scaffolds, its benefits should have been amplified for familiar characters. Instead, retrieval improved memory equally for high-knowledge and low-knowledge items, implying that prior knowledge elevates overall performance but does not interact with retrieval dynamics. This pattern converges with recent demonstrations that forgetting rates remain invariant across differences in learning strength, repetition, and age (Rivera-Lares et al., 2022, 2023, 2024). Together, these findings extend the rate-invariance phenomenon from quantitative forgetting research to the domain of semantic knowledge: the shape of forgetting—and its modification by retrieval—appears stable regardless of how well information is initially learned or semantically supported.

Our results offer a nuanced perspective on the dynamics of selective retrieval that extends prior work. Earlier evidence suggested that selective retrieval can reduce subsequent forgetting for retrieved relative to nonretrieved items, implying that retrieval may alter the rate of memory decay (Bäuml & Trißl, 2022). In the present study, however, selective retrieval produced a clear initial recall advantage for retrieved over nonretrieved items, while both item types showed comparable absolute forgetting trajectories thereafter. One possible explanation for this discrepancy concerns differences in task timing and baseline structure. In our design, selective retrieval occurred after an extended delay and was evaluated against a time-delay baseline. Under these conditions, retrieval may have reinstated the study context and elevated recall performance without changing the later temporal profile of forgetting. This interpretation is consistent with recent evidence suggesting that retrieval can shift memory accessibility while leaving the overall forgetting rate largely unchanged (Bäuml et al., 2025). More broadly, these findings indicate that the effects of selective retrieval are not uniform and may depend on specific experimental parameters, including the nature of the baseline condition, the delay structure, and the retrieval demands. Identifying these boundary conditions remains an important direction for future research. Within the present design, our results demonstrate that retrieval-induced changes in memory accessibility and forgetting dynamics operate similarly across different levels of prior knowledge.

Beyond absolute recall levels, the dynamics of relative forgetting offer a complementary perspective on memory stability. Standard accounts suggest that higher initial recall levels are associated with reduced relative forgetting (Wixted, 2022). However, complementary analyses of the power-function parameter $b$ and interval-based relative loss yielded the same general pattern, with no reliable differences in relative decay across knowledge levels. This further reinforces our primary finding that while prior knowledge elevates overall memory performance, the fundamental dynamics of forgetting and retrieval-induced stabilization remain domain general and knowledge independent.

The present results highlight selective retrieval as a robust behavioural influence on memory retention within the present paradigm. Retrieval does not eliminate forgetting but appears to reinstate and shift the forgetting trajectory by reactivating temporal context. While this effect is likely

time-limited and may diminish over extended retention intervals (Bäuml et al., 2025), its consistency across knowledge levels suggests that selective retrieval operates as a robust mechanism within the retention window examined here. Accordingly, our findings are best interpreted as evidence for a rate-invariance phenomenon within this paradigm, rather than a universal principle of memory.

The parallel forgetting functions across knowledge levels are consistent with the possibility that temporal context reactivation contributed to retrieval-induced enhancement. According to the Context Maintenance and Retrieval (CMR) model (Polyn et al., 2009; Polyn & Woodman, 2025), recalling an item reinstates the temporal context of the study episode, thereby increasing the accessibility of associated traces. Because this process is content-agnostic, it naturally explains why retrieval boosts both familiar and novel materials. Consistently, our data suggests that semantic scaffolds contribute additively—raising baseline accuracy—but do not shape the temporal dynamics of retrieval-based learning. Retrieval thus may operate through a mechanism that transcends representational richness, reorganizing memory over time through context reinstatement.

To situate our findings within the broader literature, it is important to consider both inhibitory and facilitatory effects of selective retrieval. Classical retrieval-induced forgetting (RIF) studies demonstrate that retrieving a subset of items can impair memory for related, nonretrieved items (Anderson et al., 1994). In contrast, the present results show facilitation rather than impairment, suggesting that the outcome of selective retrieval may depend on specific task condition. One key factor may be the temporal structure of the experiment. Prior work has shown that increasing the delay between study and retrieval can alter retrieval dynamics, producing a gradual shift from inhibition to facilitation (Bäuml et al., 2025; Chan et al., 2006; Kriechbaum & Bäuml, 2023). Under longer-lag conditions, retrieval may reinstate aspects of the original study context, thereby facilitating access to both retrieved and nonretrieved items. Spreading activation may provide an additional account by increasing the accessibility of related representations (Anderson, 1983; Collins & Loftus, 1975). However, this explanation would require the net facilitatory influence of spreading activation to become stronger at longer study-retrieval lags; otherwise, it would not readily account for the transition from impairment at shorter lags to facilitation at longer lags. Rather than attributing the observed effects to a single process, the present findings are consistent with a framework in which multiple mechanisms—including inhibition, context reinstatement, and associative activation—jointly shape retrieval outcomes. Finally, how prior knowledge interacts with these processes remains an open question. Existing evidence is mixed (Garcia-Bajos & Migueles, 2009; Marful et al., 2016; Migueles & García-Bajos, 2014), and further work is needed to determine whether and how prior knowledge modulates the balance between facilitation and inhibition in selective retrieval paradigms.

The present results highlight selective retrieval as a universal phenomenon of memory stabilization. Retrieval not only prevents loss but reorganizes memory by refreshing its temporal context, a process that appears invariant across levels of prior knowledge and across writing systems. This universality implies that retrieval practice should enhance learning even for novices, providing a theoretical foundation for its wide applicability in educational and training settings. Conceptually, our findings bridge the literatures on forgetting and expertise, suggesting that forgetting reflects not passive decay but an adaptive, time-structured architecture that retrieval can temporarily modify.

Future work can extend these findings along several paths. First, examining longer retention intervals—from days to weeks—will test whether knowledge-independence persists after consolidation. Second, using semantically integrated materials could reveal when retrieval interacts with schema formation, defining the boundary conditions of the current effect. Third, combining behavioural paradigms with neuroimaging or intracranial recordings may help clarify the mechanisms underlying retrieval-induced changes in memory, including whether context reactivation produces transient or longer-lasting effects. Finally, applying selective retrieval paradigms to domains such as perceptual expertise, motor learning, or classroom education could determine whether retrieval universally stabilizes memory beyond verbal tasks.

Retrieval boosts memory and modifies forgetting for both familiar and unfamiliar materials, suggesting that temporal context reactivation can function as a domain-general mechanism of retention. By integrating the literatures on expertise and forgetting, this study clarifies how retrieval reorganizes memory across time and establishes selective retrieval as a phenomenon for stabilizing knowledge.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.3758/s13423-026-03004-4.

**Author contributions** Y.C. and Z.H. equally contributed to the current research. Y.C. and B.W. designed the experiment. J.L. collected the data. Y.C., Z.H., J.L., and B.W. analysed the data. Y.C., Z.H., G.W., B.Z., and B.W. wrote and revised the paper. All authors wrote the paper together and approved the final version of the manuscript for submission.

**Funding** This research was supported by the Natural Science Foundation of Guangdong grant (2023A1515012789) to Benchi Wang.

**Data availability** The codes and data can be accessed online (https://github.com/Seraphoenix/Selective-retrieval-).

**Code availability** The codes and data can be accessed online (https://github.com/Seraphoenix/Selective-retrieval-).

## Declarations

**Ethics approval** This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Institutional Review Board of South China Normal University (2020-3-013).

**Consent to participate** Informed consent was obtained from all individual participants included in the study.

**Consent for for publication** All participants signed informed consent regarding publishing their data.

***Author note*** Z.H. and Y.C. equally contributed to the current research. Y.C. and B.W. designed the experiment. J.L. collected the data. Y.C., Z.H., J.L., and B.W. analyzed the data. Y.C., Z.H., G.W., B.Z., and B.W. wrote and revised the paper. All authors wrote the paper together and approved the final version of the manuscript for submission. All authors declared no competing interests. This research was supported by the Natural Science Foundation of Guangdong grant (2023A1515012789) to B.W.

**Conflicts of interest/Competing interests** All authors have no relevant financial or nonfinancial interests to disclose.

## References


Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 261–295.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1063.

Bäuml, K. H. T., & Dobler, I. M. (2015). The two faces of selective memory retrieval: Recall specificity of the detrimental but not the beneficial effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(1), 246.

Bäuml, K. H. T., Meixensperger, S. R., & Hirsch, M. L. (2025). Reinstating memories' temporal context at encoding causes Sisyphus-like memory rejuvenation. *Proceedings of the National Academy of Sciences of the United States of America, 122*(32), Article e2505120122.

Bäuml, K.-H., & Schlichting, A. (2014). Memory retrieval as a self-propagating process. *Cognition, 132*(1), 16–21.

Bäuml, K.-H.T., & Trißl, L. (2022). Selective memory retrieval can revive forgotten memories. *Proceedings of the National Academy of Sciences of the United States of America, 119*(8), Article e2114377119. https://doi.org/10.1073/pnas.2114377119

Chan, J. C., McDermott, K. B., & Roediger, H. L., III. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135*(4), 553.

Chen, B., & Peng, D. (1998). The influence of concreteness of words on word recognition. *Acta Psychologica Sinica, 30*(4), 387–393.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407.

Dictionary Editing Office, Institute of Linguistics, Chinese Academy of Social Sciences. (2016). *Xiandai Hanyu Cidian [A dictionary of current Chinese]* (7th ed.). Commercial.

Ding, G., Taft, M., & Zhu, X. (2000). The representation of radicals that can be used as characters. *Acta Psychologica Sinica, 32*(suppl), 21–26.

Ebbinghaus, H. (1885). *Über das gedächtnis: Untersuchungen zur experimentellen psychologie [On memory: Studies in experimental psychology]*. Duncker & Humblot.

Ferrand, L., & Grainger, J. (2003). Homophone interference effects in visual word recognition. *Quarterly Journal of Experimental Psychology, 56*(3), 403–419.

Garcia-Bajos, E., & Migueles, M. (2009). When stereotype knowledge prevents retrieval-induced forgetting. *Acta Psychologica, 131*(1), 63–71.

Kliegl, O., Carls, T., & Bäuml, K. H. T. (2019). How delay influences search processes at test. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(12), 2174.

Kriechbaum, V. M., & Bäuml, K. H. T. (2023). The critical importance of timing of retrieval practice for the fate of nonretrieved memories. *Scientific Reports, 13*(1), 6128.

Marful, A., Gómez-Ariza, C. J., Barbon, A., & Bajo, T. (2016). Forgetting "novel" but not "dragon": The role of age of acquisition on intentional and incidental forgetting. *PLoS ONE, 11*(5), Article e0155110.

Migueles, M., & García-Bajos, E. (2014). The role of exemplar typicality and encoding strategies in category retrieval-induced forgetting. *Memory, 22*(3), 212–221.

Nickl, A. T., & Bäuml, K. H. T. (2024). To-be-forgotten information shows more relative forgetting over time than to-be-remembered information. *Psychonomic Bulletin & Review, 31*(1), 156–165.

Perfetti, C. A., & Liu, Y. (2005). Orthography to phonology and meaning: Comparisons across and within writing systems. *Reading and Writing, 18*(3), 193–210. https://doi.org/10.1007/s11145-004-2344-y

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116*(1), 129–156. https://doi.org/10.1037/a0014420

Polyn, S. M., & Woodman, G. F. (2025). Capacity not required: A long-term memory model that exhibits key signatures of working memory. *Psychological Review*. https://doi.org/10.1037/rev0000588

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X: Control of Language Processes, 32*, 531–556.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95*(3), 318.

Rivera-Lares, K., Baddeley, A., & Della Sala, S. (2024). Influence of degree of learning on rate of forgetting of tonal sequences. *Memory & Cognition*. https://doi.org/10.3758/s13421-024-01597-6

Rivera-Lares, K., Logie, R., Baddeley, A., & Della Sala, S. (2022). Rate of forgetting is independent of initial degree of learning. *Memory & Cognition, 50*(8), 1706–1718. https://doi.org/10.3758/s13421-021-01271-1

Rivera-Lares, K., Sala, S. D., Baddeley, A., & Logie, R. (2023). Rate of forgetting is independent from initial degree of learning across different age groups. *Quarterly Journal of Experimental Psychology, 76*(7), 1672–1682. https://doi.org/10.1177/17470218221128780

Roediger, H. L., III., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255.

Sala, G., & Gobet, F. (2017). Experts' memory superiority for domain-specific random material generalizes across fields of expertise: A meta-analysis. *Memory & Cognition, 45*(2), 183–193. https://doi.org/10.3758/s13421-016-0663-2

Taft, M., Zhu, X., & Ding, G. (2000). The relationship between character and radical representation in Chinese. *Acta Psychologica Sinica*, *32*(Suppl.), 3–12.

Wang, B., & Theeuwes, J. (2018). Statistical regularities modulate attentional capture independent of search strategy. *Attention, Perception, & Psychophysics, 80*(7), 1763–1774.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131*(1), 48.

Wing, E. A., Burles, F., Ryan, J. D., & Gilboa, A. (2022). The structure of prior knowledge enhances memory in experts by reducing interference. *Proceedings of the National Academy of Sciences of the United States of America, 119*(26), Article Article e2204172119. https://doi.org/10.1073/pnas.2204172119

Wixted, J. T. (2022). Absolute versus relative forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(12), 1775–1786.

Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition, 25*(5), 731–739. https://doi.org/10.3758/BF03211316

Yeh, S. L., & Li, J. L. (2002). Role of structure and component in judgments of visual similarity of Chinese characters. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 933.

Zhang, J., Du, J., & Dai, L. (2020). Radical analysis network for learning hierarchies of Chinese characters. *Pattern Recognition, 103*, Article 107305.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.