Acta Psychologica 255 (2025) 104901

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

The superior memory effect of insightful learning: The role of associative novelty

Zhonglu Zhang [a,*], Yuye Si [a], Lan Wang [a], Kaiyun He [a,b,*]

[a] Department of Psychology, School of Education, Guangzhou University, Guangzhou 510006, China
[b] Institute of Psychogerontology, School of Psychology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberg, Germany

ARTICLE INFO

Keywords:
Insightful learning
Associative memory
Metamemory
Novelty
Chinese logogriph

ABSTRACT

Previous research has confirmed the advantage of insightful learning in item memory, but its impact on associative memory, particularly the role of novelty, remains unclear. This study aims to investigate the role of novelty in insightful learning on associative memory and metamemory using a Chinese logogriph learning-testing paradigm. Participants were asked to judge whether the "Midi" word/character (the answer) matched the "Mimian" phrase (the problem) in both the high- and low-novelty conditions. Then they took immediate (2 min later) and delayed (24 h later) recognition tests to identify whether the "Mimian"-"Midi" (phrase-character) pairs were previously presented (old) or not (new), and assessed their retrospective confidences. Results showed that individuals had higher recognition rates for high- (vs. low-) novelty logogriphs, and higher retrospective confidence for recognized pairs. This research indicates that novelty enhances both associative memory and the confidence judgment of metamemory monitoring in insightful learning.

1. Introduction

Traditional education tends to focus on memorization and comprehension through repeated presentations, which can be inefficient and time-consuming (e.g., Chrabaszcz et al., 2023; Crouse, 2024). Conversely, insightful learning, a relatively neglected approach in education, can achieve a "learning in one go" effect (Ludmer et al., 2011). Insight refers to individuals suddenly discovering or understanding truths or answers that were previously unknown, experiencing an "aha!" moment of "Oh! That's how it is!" (Bowden & Jung-Beeman, 2007). Since Köhler introduced the concept of "insight", he viewed it as a learning mechanism where learning is achieved through sudden realization rather than through trial and error (Köhler, 1925). Nevertheless, it wasn't until the late 1970s that individuals started to take notice of and slowly appreciate the influence of insight on learning and memory.

Prior research has primarily induced insight by having individuals attempt to answer questions before presenting the answers in order to explore its impact on memory. Studies have found that individuals exhibit better memory performance for item information under insightful learning conditions compared to conventional (non-insightful) learning methods (Danek & Wiley, 2024; Lai et al., 2019). The insight memory advantage effect has robustness across task types (verbal tasks or picture tasks), memory modes (recall or recognition), levels of consciousness (conscious or unconscious memory), age groups (young or old), and time spans (minutes or days) (Auble et al., 1979; Kizilirmak et al., 2016, 2021; Shen et al., 2021, 2022; Wills et al., 2000; Zhang & Liu, 2021). Specifically, Auble et al. (1979) found that compared to presenting verbal logographs and answer hints simultaneously or giving incorrect hints (non-insight conditions), people could better remember the sentence when the answer/hint was presented 5 s later (insight conditions). Similarly, Wills et al. (2000) had participants sketch dot patterns (insight condition) or complete patterns (non-insight condition), and then presented real pictures 5 s later. The results showed that participants had better memory for the pictures in the dot pattern condition.

Kizilirmak and colleagues compared sudden understanding (insight) or continued lack of understanding (non-insight) conditions based on a remote association task, and found that both young and old participants had a higher recognition rate in the sudden understanding condition (compared to not understanding) (Kizilirmak et al., 2016, 2021). Shen et al. (2021) had participants watch (conventional or creative) advertisements and report whether they had an insight experience, and conducted recall and recognition tests 5 min or 3 days later. The findings indicated that individuals who reported having insight had higher recall

* Corresponding authors.
E-mail addresses: zzhang@gzhu.edu.cn (Z. Zhang), kaiyun.he@fau.de (K. He).

https://doi.org/10.1016/j.actpsy.2025.104901
Received 25 November 2024; Received in revised form 24 February 2025; Accepted 10 March 2025
Available online 15 March 2025
0001-6918/

and recognition rates for pictures compared to those without insight (Shen et al., 2021), as well as better unconscious memory effects (Shen et al., 2022). Zhang and Liu (2021) used an appreciation-recognition paradigm to induce insight by providing key words of the poem, revealing the insight advantage effect and mechanism for the entire poem. Specifically, participants with insight had faster recognition responses than those without insight; at the neural level, neither new poetic lines nor non-insight conditions showed early FN400 effects compared to the insight condition. Only between 550 and 750 ms, the insight condition showed topographical old-new effects. Furthermore, Ludmer et al. (2011) had participants first attempt to identify blurry pictures and then presented the answer to induce perceptual insight, finding that individuals were more confident in correctly recognizing pictures one week later compared to incorrectly recognized pictures. This indicates that induced insight can also improve confidence judgments of metamemory.

As a typical form of creative thinking, the nature of insight is establishing novel associations between elements in a problem context (Bowden & Jung-Beeman, 2007; Chen et al., 2021; Huang et al., 2018; Zhao et al., 2017, 2024). Novelty or novel associations may be the key factor leading to the insight memory advantage effect (Chen et al., 2021; Zhao et al., 2024). Previous research has provided direct or indirect evidence for this. Zhao and colleagues distinguished between novel and conventional answers, requiring participants to select matching answers to idioms from mixed items. They revealed that people had a better recognition of novel answers in immediate and delayed tests, and the advantage of novel semantic associations was mainly related to activation in the right hippocampus (Chen et al., 2021), as well as longer fixation time (Zhao et al., 2024). Consistently, research has also found that people have higher recognition and confidence (mate-memory) for creative advertisements compared to conventional advertisements (Shen et al., 2021, 2022). Wu et al. (2018) observed that in both explicit and implicit recognition tasks of Chinese character decomposition, novel and appropriate chunking had better memory effects compared to other conditions (novel and inappropriate, familiar and appropriate, familiar and inappropriate). Moreover, Ding et al. (2023, 2024) found that subjective ratings of novelty positively predicted individuals' recall and recognition performance for multi-purpose answers in an alternative uses task. Furthermore, Ding et al. (2024) found that activations in the hippocampus predicted better recognition.

In summary, previous research has consistently found that insightful learning has an advantage in item memory (such as answers) and has revealed the role of novelty in this process. However, there has been limited research on the impact of insight on associative memory, especially the role of novelty in it. Associative memory refers to memory for two or more units of information that were presented together in a prior event (e.g., Bader et al., 2010; Mohanty & Naveh–Benjamin, 2018; Naveh-Benjamin, 2000). Associative memory not only stores information about items or events themselves, but more importantly, it stores information about the relationships between items (Buchler et al., 2008; Zhao & Guo, 2023). Studies have shown that item memory and associative memory have shared but distinguishable mechanisms: at the cognitive level, both involve a recollection process. However, item memory, not associative memory, requires the involvement of familiarity (Yonelinas, 1997). Neuroimaging studies have found that at the neural level, both item memory and associative memory rely on the medial temporal lobe. However, item memory mainly relies on the perirhinal cortex, while associative memory mainly relies on the hippocampus (Davachi, 2006; Diana et al., 2010). Additionally, research has found that encoding items before associations can impair associative memory by consuming encoding resources, but it can also promote associative memory through inter-item associations (Forester & Kamp, 2023). Obviously, item memory and associative memory are distinct yet related. It is speculated that the insight novelty effect found in item memory may only be "the tip of the iceberg" (e.g., Chen et al., 2021), and the more fundamental aspect may be that insight novelty promotes associative memory (between items). Furthermore, the EEG findings of Zhang and Liu (2021) on the insight memory advantage in poetry appreciation also suggest that insight mainly excels in recollection and integrating semantic associations of poems (as evidenced by significant P550–750 old/new effects in the parietal region) rather than processing item familiarity (no significant FN400 old/new effects). Considering that the essence of insight lies in establishing novel associations, meaning new relationships between items (Bowden & Jung-Beeman, 2007; Chen et al., 2021), this study hypothesizes that novelty can enhance associative memory.

In addition, unlike item or associative memory, metamemory refers to an individual's monitoring of these memory activities (Koriat & Goldsmith, 1996; Nelson & Narens, 1990), which can be measured through retrospective confidence judgments (Chua et al., 2009; Dougherty et al., 2005; Hu et al., 2022). Previous studies have indicated that metamemory monitoring plays an important role in regulating individuals' learning activities, such as deciding what to restudy or what not (Robey et al., 2017, 2022). In this regard, it is interesting and valuable to study the influence of insight and novelty on metamemory monitoring. To date, two studies have explored and found that creative insight can enhance metamemory monitoring, as indexed by retrospective confidence ratings (Ludmer et al., 2011; Shen et al., 2021). The influence of insight, especially its novelty component, on association-related metamemory monitoring still needs to be investigated.

The study aims to explore the impact of induced insight on associative memory and its metamemory, particularly examining the role of novelty in it. For this purpose, we employed the Chinese logographs task, which has been traditionally used in studying insight and creativity (Chen et al., 2021; Huang et al., 2018; Mai et al., 2004; Qiu et al., 2006; Zhang et al., 2023; Zhao et al., 2017, 2024). Specifically, a logograph is normally a phrase-word pair consisting of a "Mimian" phrase (similar to a problem) and a "Midi" word/character (similar to an answer). Previous studies have indicated that once people get or comprehend the solution (the "Midi" character) according to the information conveyed by the problem (the "Mimian" phrase), they would probably have "aha!" emotional experiences (e.g., Chen et al., 2021; Mai et al., 2004; Zhang et al., 2023; Zhao et al., 2017), a golden marker of creative insight (Bowden & Jung-Beeman, 2007). Moreover, this insight moment is achieved mainly through the formation of novel and appropriate associations (Bowden et al., 2005; Chen et al., 2021; Zhao et al., 2017). For example, in the phrase-word pair ["六十天" (reads: liu shi tian)-"朋"(reads: peng)], the three-word "Mimian" phrase "六十天" means "sixty days". The "Midi" word/character "朋" is made up of two characters "月"(reads: yue),which means month. Thus, people can get or comprehend the solution (the "Midi" word/character "朋") according to the information conveyed by the problem (the "Mimian" phrase "六十天"), as they share appropriate associations (Generally, two months are approximately equal to sixty days).

Moreover, the experiment would manipulate Chinese logographs with high and low levels of novelty, which would be distinguished by categorizing the scores of novelty ratings on a 5-point scale (e.g., the top 30 % of the rating scores would be categorized as high-novelty logographs and the bottom 30 % of the rating scores would be categorized as low-novelty logographs). Then, we would compare the differences between the two conditions in the associative memory of the "Mimian-Midi" pair (a phrase-word pair) and its metamemory in both immediate and delayed recognition tests. The study primarily hypothesizes that, compared to the low-novelty condition, individuals in the high-novelty condition would have better memory effects on the "Mimian-Midi" associations. Meanwhile, we also examine the influence of novelty on the retrospective confidence rating of metamemory monitoring.

2. Methods

2.1. Experimental design

The experiment aims to investigate the impact of novelty on the "Mimian"-"Midi" associative memory. A within-subject design of 2 (logogriph novelty: high novelty vs. low novelty) × 2 (recognition time: immediate vs. delayed) was used, with the dependent variables being comprehension rate, comprehension response time, "aha!" emotional experience during the learning phase, as well as associative recognition rate, hit rate, response time, and confidence value of the "Mimian"-"Midi" pairs during the recognition phase.

2.2. Participants

Using the More-power6.4 software to estimate sample size (Campbell & Thompson, 2012), based on previous research (Zhou, 2022), the effect size on associative memory performance was about 0.22. With a power of 0.8 and an α level of 0.05, it was calculated that a sample size of 30 participants was needed. Forty-nine university students were recruited for the experiment, out of which six participants were excluded because there was more than a 50 % probability that the participant judged "match" on the 12 non-matching logographs (the "Mimian" phrases did not match the "Midi" words/characters), or there was less than a 50 % probability that the participants judged "non-match" on the 48 matching logogriphs (the "Mimian" phrases matched the "Midi" characters). This remained a total of 43 valid participants (24 females, 19 males, $M_{age} = 20.26 \pm 1.56$ years). Two participants were left-handed, while the rest were right-handed with normal or corrected-to-normal vision. Participants provided informed consent before the experiment and received compensation afterwards. The study was approved by the local ethics committee.

2.3. Materials

179 sets of Chinese logogriphs were initially selected from books and online sources, with each logogriph pair consisting of a three-character "Mimian" phrase (e.g., "六十天") and a "Midi" character (e.g., "朋"). Then 20 university students were asked to evaluate the dimensions of comprehensibility, appropriateness and novelty of these logogriphs on a 5-point Likert scale (1 = not at all, 5 = very). The inter-rater reliability of the dimensions for scoring was evaluated by the intraclass correlation coefficient (ICC) (comprehensibility: 0.89; appropriateness: 0.87; novelty: 0.56). Logogriphs with average ratings of 2.5 or lower in both comprehensibility and appropriateness were eliminated, leaving 169 logogriphs. These were then ranked by novelty, with 58 logogriphs deemed highly novel (top 30 %) and 57 deemed low in novelty (bottom 30 %).

Considering the impact of imagery and concreteness on memory (Li et al., 2020), an additional 30 students were asked to rate the familiarity, imagery, and concreteness of the remaining logogriphs on a 7-point Likert scale (1 = not at all, 7 = very). The inter-rater reliability of the dimensions for scoring was as follows (ICC): familiarity: 0.96; imagery: 0.90; concreteness: 0.94. Controlling for factors of comprehensibility, appropriateness, familiarity, concreteness, imagery, word frequency, and stroke count, the final 48 logogriphs were selected for the formal experiment, with 24 high-novelty and 24 low-novelty logogriphs (see Table 1).

These 48 logogriphs were first used for comprehension judgment in the learning phase. For all 48 logogriphs, the "Midi" character was the answer/solution to the problem (the "Mimian" phrase). In other words, they matched each other in a way that the "Midi" character (e.g., "朋") and the "Mimian" phrase (e.g., "六十天") had appropriate associations [two months are approximately equal to sixty day]. This provided the chance for people to comprehend the solution (the "Midi" character) according to the information conveyed by the problem (the "Mimian" phrase). However, there was a possibility that participants might press the button indicating "match" all the time if they realized all the 48 "Mimian"-"Midi" pairs matched each other. For this reason,12 additional logogriph pairs were included in the learning phase where the "Mimian" phrase and "Midi" phrase were mismatched. In other words, a "Midi" word/character of one logogriph was randomly paired with a "Mimian" phrase of another logogriph. These 12 non-matching pairs served as catch trials, which not only helped to detect whether participants were engaged in the experiment seriously (the data of the participants who indicated a "match" judgment on the 12 non-matching pairs would be excluded), but also reduced response tendency.

Table 1
Different dimensions of the Chinese logogriphs ($M \pm SD$).

Index	High-novelty (N = 24)	Low-novelty (N = 24)	t	P
Comprehensibility	4.10 ± 0.36	4.24 ± 0.44	−1.15	0.26
Appropriateness	3.95 ± 0.37	3.95 ± 0.37	−1.25	0.22
Novelty	3.93 ± 0.14	3.22 ± 0.11	19.56	0.00
Familiarity	4.52 ± 1.31	4.91 ± 1.09	−1.13	0.27
Concreteness	4.47 ± 1.30	4.61 ± 1.07	−0.43	0.67
Imagery	4.23 ± 1.11	4.42 ± 0.78	−0.70	0.49
Strokes	8.21 ± 3.53	6.79 ± 2.89	1.52	0.14
Frequency	0.03 ± 0.10	0.04 ± 0.12	−0.66	0.33

In the testing phase, half of the 24 high-novelty logogriphs and half of the 24 low-novelty logogriphs were randomly pooled into the immediate test and the delayed test, respectively. In this way, any phrase-word pair was tested only once so that the influence of testing effects was excluded (Bacso & Marmurek, 2016; Krieglstein et al., 2024). Moreover, in order to measure and compute the associative recognition rate (see details in Section 2.5: Data Analysis), the original phrase-word pairs ("Mimian"-"Midi") were further equally divided into two groups, with half being presented as the same/old pair as they did in the learning phase, and the other half being presented as rearranged/new pairs, which were different from those in the learning phase for both the high-novelty and low-novelty conditions. Ultimately, each testing phase (immediate or delayed) included 6 old phrase-word pairs of high-novelty logogriphs together with another 6 rearranged pairs as counterparts, and 6 old phrase-word pairs of low-novelty logogriphs together with another 6 rearranged pairs as counterparts. For example, the phrase-word pair (e.g., "六十天"-"朋") was first presented in the learning phase, and the same/old pair (e.g., "六十天"-"朋") or the rearranged/new phrase-word pair (e.g., "六十天"-"做") could be presented for recognition judgment in the testing phase. In this way, participants could judge whether the presented phrase-word pairs were old (appeared together in the learning phase) or new/rearranged (not appeared together in the learning phase).

It should be noted that the phrase "Mimian" and the word "Midi" matched each other in the original pair, but did not match in the rearranged pair. There was a possibility that participants might have used "matching" strategies to make their recognition judgments. In other words, their final recognition scores probably included both their "real memory" and their "matching strategies" scores. To reduce and measure this potential influence, 48 new logograph pairs were included in the recognition testing, with 24 matching pairs and 24 non-matching pairs. This allowed for the measurement and exclusion of the potential influence of "matching" strategies on memory performance by comparing the difference in recognition scores between the new "matching" and "non-matching" pairs. In total, 96 "Mimian-Midi" (phrase-word) pairs were randomly presented during the recognition phases, including both the immediate and delayed phases.

2.4. Procedure

The experiment was conducted using E-prime 2.0 and consisted of four stages: learning, interference, immediate test, and delayed test (as shown in Fig. 1). In the learning stage, a fixation point "+" was presented on the screen for 500 milliseconds, followed by a three-character "Mimian" phrase (the problem of a logograph) for 3 s. Participants were required to press the "ENTER" key if they thought of the answer, otherwise they did not press any key. The "Mimian"-"Midi" (phrase-word) pair of the logograph was then presented, and participants had 5 s to judge whether the "Midi" word/character matched (or was the answer of) the three-character phrase (the problem) of the logograph (see Fig. 1). They would judge "match" if they understood the appropriate associations between the "Midi" word/character and the three-character "Mimian" phrase. Otherwise, they would judge "non-match". Participants then rated their "aha!" experience on a scale of 1 to 3, with 1 indicating no "aha!" experience, 2 indicating neutral, and 3 indicating a strong "aha!" experience (Bowden & Jung-Beeman, 2007). The program then automatically moved on to the next trial. Trials were presented randomly, with participants taking a break every 24 trials.

Based on previous research (Shen et al., 2021), an interference stage was introduced where participants had to perform a task of subtracting 3 continuously and write down the results on paper, for example: 497, 494… This stage lasted for 3 min.

In the immediate recognition stage, participants had 3 s to judge whether the "Mimian"-"Midi" pair had appeared together in the learning stage. They pressed the "F" key if they had appeared together, and the "J" key if they had not. Key presses were balanced across participants. Confidence in their recognition judgment was then rated on a scale of 1–5: 1 indicating completely not confident, 2 indicating not confident, 3 indicating neutral, 4 indicating confident, and 5 indicating very confident.

In the delayed testing stage, participants underwent another testing task with the other half of the materials 24 h later. Prior to the formal experiment, participants completed several practice sessions to familiarize themselves with the procedures and tasks.

2.5. Data analysis

Statistical analysis was conducted on the comprehension rate and recognition performance under different conditions as follows. 1) Comprehension rate: (number of correctly understood logographs) / [total number of logographs ("Mimian-Midi") pairs] – (number of solved logographs during 3 s). Here the very few number of solved logographs were excluded from further analysis, as this study mainly investigated the effect of induced insight (Kizilirmak et al., 2016; Ludmer et al., 2011). 2) Hit rate: (number of old logographs pairs recognized as old) / (sum of old logographs pairs recognized as old or new); 3) False alarm rate: (number of rearranged "Mimian"-"Midi" pairs judged as old) / (sum of rearranged "Mimian"-"Midi" pairs judged as old or new) 4) Associative recognition rate: P (hit rate) - P (false alarm rate) (Mohanty & Naveh–Benjamin, 2018; Zhao & Guo, 2023); 5) The recognition time and confidence level were calculated for the hit trials. 6) To detect whether participants used matching strategies for recognition judgments (making the recognition judgment according to whether the "Mimian" matched the "Midi" in the pairs), we measured the probability of people judging "old" for the newly-introduced matching and non-matching "Mimian"-"Midi" pairs, as well as their difference ($P_{matching\_strategies\_effect}$). 7) Finally, we computed a modified associative recognition rate by following this formula: P (hit rate − false alarm rate) − $P_{matching\_strategies\_effect}$.

The paired sample *t*-tests were performed on the comprehension rate, response time, and "aha!" experience for high and low novelty conditions in the learning stage. The 2 × 2 repeated measures ANOVAs were conducted on the hit rates, false alarm rates, (modified) associative recognition rates, recognition response time, and confidence levels in the recognition stage. Furthermore, logarithmic transformation was applied to the reaction times for both the comprehension and recognition tasks prior to analysis, ensuring that the overall data distribution more closely approximated a normal distribution.

3. Results

3.1. Learning stage

A paired samples t-test was conducted to compare the solution rate, the comprehension rate, reaction time and "aha!" experience value between high- and low-novelty logographs (see Table 2). The findings indicated that there was a significant difference of the solution rates between the two conditions, t (42) = 4.60, $p < .001$, Cohen's $d = 0.70$. There was no significant difference in the comprehension rate between the two conditions, t (42) = 1.70, $p = .096$, Cohen's $d = 0.26$. Moreover, the reaction time(logarithmic) was significantly longer in the high-novelty group compared to the low-novelty group, t (42) = 2.19, p =

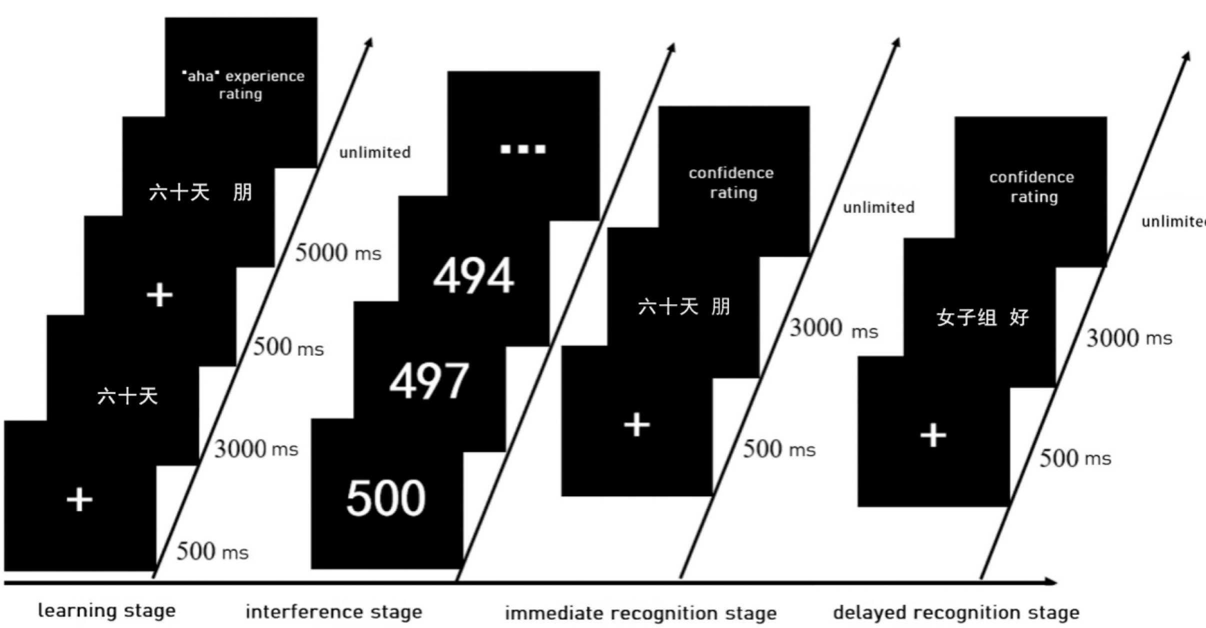

Fig. 1. An example of the experimental procedure.

Table 2
Solution rate, Comprehension rate, Response time and “aha!” experience in the learning stage ($M \pm SD$).

Logogriph type	Solution rate	Comprehension rate	Response time (ms)	“aha!” Experience
High novelty	0.08 ± 0.13	0.85 ± 0.09	2325.03 ± 437.25	2.79 ± 0.17
Low novelty	0.17 ± 0.20	0.82 ± 0.11	2223.39 ± 449.49	2.63 ± 0.28

.03, Cohen’s $d = 0.33$. The “aha!” experience was more intense in the high- (vs. low-) novelty group, t (42) = 4.59, $p < .001$, Cohen’s $d = 0.70$.

3.2. Recognition stage

The descriptive statistics results are shown in Table 3 and Fig. 2. A repeated measures analysis of variance was conducted separately for the hit rate of old logographs, (logarithmic) reaction time for hits, confidence scores, ratings, (modified) associative recognition rates, and false alarm rates for new logographs, with a 2 (novelty of logographs: high vs. low) × 2 (recognition time: immediate vs. delayed) design.

The main effect of novelty of logographs on hit rate was not significant, $F\,(1, 42) = 2.56$, $p = .12$, $\eta_p^2 = 0.06$. However, recognition time did significantly impact hit rate, with immediate recognition yielding a higher rate compared to delayed recognition, $F\,(1, 42) = 6.76$, $p = .01$, $\eta_p^2 = 0.14$. An interaction effect between logograph novelty and recognition time was observed, $F\,(1, 42) = 4.42$, $p = .04$, $\eta_p^2 = 0.10$. Further analysis showed that there was no significant difference in hit rate between the high-novelty group (M = 0.95, SD = 0.11) and the low-novelty groups (M = 0.95, SD = 0.13) under immediate recognition, t = 0.21, p = .84, Cohen’s d = 0.03, 95 % CI = [−0.05, 0.04]. However, under delayed recognition, the high-novelty group (M = 0.92, SD = 0.15) exhibited a significantly higher hit rate than the low-novelty group (M = 0.85, SD = 0.13), t = 2.16, p = .04, Cohen’s d = 0.33, 95 % CI = [0.01, 0.15]. Moreover, for low-novelty logographs, the hit rate was significantly higher in immediate tests (M = 0.95, SD = 0.13) compared to delayed tests (M = 0.85, SD = 0.13) for the low-novelty logographs, t = 2.91, p = .01, Cohen’s d = 0.44, 95 % CI = [0.03, 0.18]. Conversely, no difference in hit rates was found between immediate and delayed recognition for high-novelty logographs (immediate: M = 0.95, SD = 0.11; delayed: M = 0.92, SD = 0.15; t = 0.10, p = .33, Cohen’s d = 0.15, 95 % CI = [−0.03, 0.08]).

For response time (logarithmic), the main effects of novelty, time and their interaction were not significant, [novelty: $F\,(1, 42) = 0.23$, $p = .63$, $\eta_p^2 = 0.00$]; [recognition time: $F\,(1, 42) = 0.17$, $p = .68$, $\eta_p^2 = 0.00$]; [interaction: $F\,(1, 42) = 0.92$, $p = .34$, $\eta_p^2 = 0.02$].

There was no significant main effect of novelty [$F\,(1, 42) = 0.64$, $p = .43$, $\eta_p^2 = 0.02$] or recognition time [$F\,(1, 42) = 0.77$, $p = .39$, $\eta_p^2 = 0.02$]. However, the interaction effect of novelty and recognition time on associative recognition rate was significant, $F\,(1, 42) = 6.09$, $p = .02$, $\eta_p^2 = 0.13$. Further analysis revealed that there was no significant difference in associative recognition rate between high-novelty (M = 0.74, SD = 0.27) and low- novelty (M = 0.78, SD = 0.28) logographs in the immediate test (t = 1.23, p = .22, Cohen’s d = 0.19, 95 % CI = [−0.12, 0.03]); In delayed recognition, the high-novelty group (M = 0.77, SD = 0.26) had a significantly higher associative recognition rate compared to the low-novelty group (M = 0.68, SD = 0.29), with t = 2.56, p = .01, Cohen’s d = 0.39, and 95 % CI = [0.02, 0.20]. Immediate recognition showed a higher recognition rate than delayed recognition under low-novelty conditions (Immediate: M = 0.78, SD = 0.28; delayed: M = 0.68, SD = 0.29), with t = 2.39 p = .02, Cohen’s d = 0.12, 95 % CI = [0.02, 0.22]. Recognition rates between them did not differ under high-novelty conditions (immediate: M = 0.74, SD = 0.27; delayed: M = 0.77, SD = 0.26; t = 0.76, p = .44, Cohen’s d = 0.12, 95 % CI = [−0.12, 0.06]).

Significant differences were found in confidence ratings for correctly recognized old matching logographs based on novelty levels, indicating higher confidence for high-novelty logographs compared to low-novelty ones [$F(1, 42) = 11.85$, $p = .001$, $\eta_p^2 = 0.22$]. Additionally, recognition time influenced confidence ratings, showing higher ratings for immediate recognition over delayed recognition [$F(1, 42) = 6.70$, $p = .01$, $\eta_p^2 = 0.14$]. However, the interaction effect between novelty and recognition time was not significant [$F(1, 42) = 0.61$, $p = .44$, $\eta_p^2 = 0.01$].

Regarding false alarm rates for new logographs, the main effects of logograph novelty and recognition time were significant, with higher false alarm rates for a new matching “Mimian”-“Midi” pair compared to a new non-matching “Mimian”-“Midi” pair, $F(1, 42) = 28.63$, $p < .001$, $\eta_p^2 = 0.41$, and higher false alarm rates for immediate compared to delayed recognition, $F(1, 42) = 6.06$, $p = .02$, $\eta_p^2 = 0.13$. The interaction effect of logograph novelty and recognition time was not significant, $F(1, 42) = 1.62$, $p = .21$, $\eta_p^2 = 0.04$, indicating a potential confounding effect of “matching” strategies.

To exclude the influence of “matching” strategies on associative recognition, a modified associative recognition rate formula was used. Results showed that the main effects of logograph novelty and recognition time were not significant, $F\,(1, 42) = 0.37$, $p = .544$, $\eta_p^2 = 0.01$; $F\,(1, 42) = 2.36$, $p = .132$, $\eta_p^2 = 0.05$. In contrast, there was a significant interaction of novelty and recognition time on the modified associative recognition rate, $F(1, 42) = 7.45$, $p = .009$, $\eta_p^2 = 0.15$. Further analysis revealed no significant difference in modified associative recognition rates between high (M = 0.65, SD = 0.32) and low (M = 0.70, SD = 0.29) novelty logographs in immediate recognition (t = −1.53, p = .133, Cohen’s d = 0.23, 95 % CI = [−0.13, 0.02]), but a significantly higher rate for the high-novelty group (M = 0.65, SD = 0.29) than the low-novelty group (M = 0.56, SD = 0.30) in delayed recognition, t = 2.63, p = .01, Cohen’s d = 0.23., 95 % CI = [0.03, 0.20]. Under conditions of low novelty, modified associative recognition rates were significantly higher for the immediate test compared to the delayed test (immediate: M = 0.70, SD = 0.29; delayed: M = 0.56, SD = 0.30), t = 2.99, p = .005, Cohen’s d = 0.46, 95 % CI = [0.06, 0.28]. However, under the high-novelty condition, there was no significant difference between them (immediate: M = 0.65, SD = 0.32; delayed: M = 0.65, SD = 0.29; t = 0.07 p = .95, Cohen’s d = 0.01, 95 % CI = [−0.11, 0.11]).

4. Discussion

This study replicated the decay law of memory over time (Wixted, 2004). Participants displayed increased associative recognition rates and confidence levels during the immediate test as opposed to the delayed test. This observation suggests a reduction in recognition performance and monitoring of associative information with longer time

Table 3
Hit rates, reaction times, (modified) associative recognition rates and confidence values under different conditions ($M \pm SD$).

Recognition time	Novelty of logograph	Hit rates	Reaction times (ms)	Associative recognition rate	(modified) Associative recognition rate	Confidence ratings
Immediate	High novelty	0.95 ± 0.11	1285.68 ± 243.38	0.74 ± 0.27	0.65 ± 0.32	4.92 ± 0.24
	Low novelty	0.95 ± 0.13	1364.42 ± 414.94	0.78 ± 0.28	0.70 ± 0.29	4.71 ± 0.56
Delayed	High novelty	0.92 ± 0.15	1438.92 ± 322.50	0.77 ± 0.26	0.65 ± 0.29	4.80 ± 0.40
	Low novelty	0.85 ± 0.21	1468.57 ± 456.67	0.68 ± 0.29	0.56 ± 0.30	4.48 ± 0.91

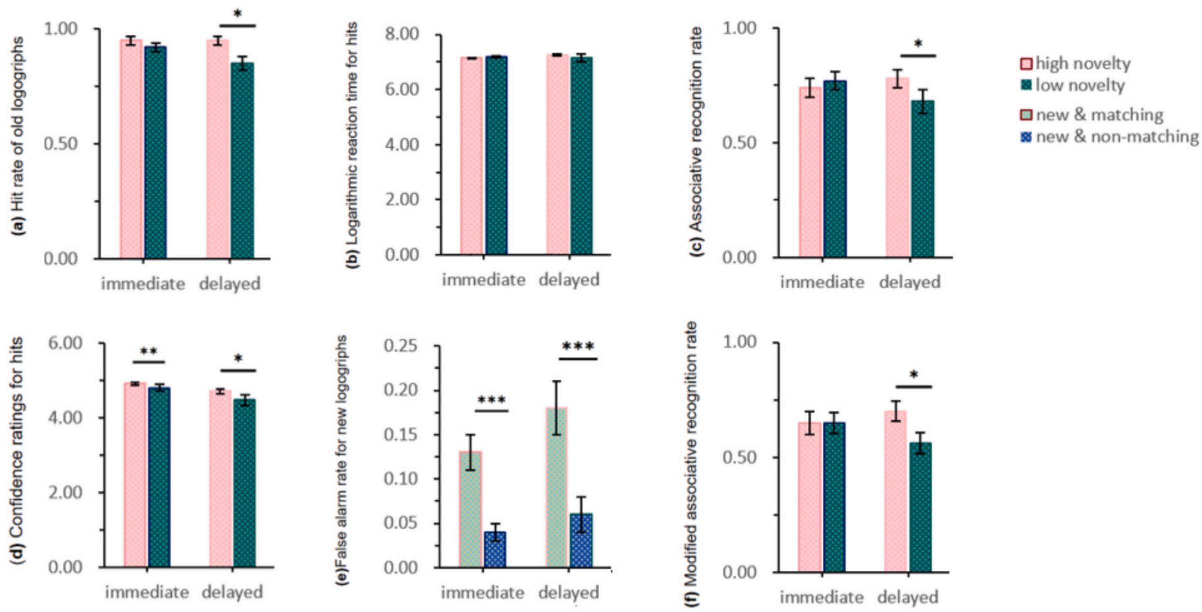

Fig. 2. Behavioral results under different conditions. (a) Hit rate of old logographs; (b) Logarithmic reaction time for hits; (c) Associative recognition rate; (d) Confidence ratings for hits; (e)False alarm rate for new logographs; (f) Modified associative recognition rate (Note: Error bars represent standard error; *$p < .05$, **$p < .01$, ***$p < .001$).

intervals, indirectly validating the experimental results.

We extended the memory advantage effect of insight on associative memory. Item memory and associative memory are components of episodic memory but vary in terms of concepts and mechanisms. Item memory involves memory representations of individual items, such as answers, primarily achieved through joint familiarity and recollection processes. Associative memory, on the other hand, focuses on memory representations of relationships between items, like phrase-word pairs, primarily facilitated by recollection processes (Buchler et al., 2008; Yonelinas et al., 2010; Zhao & Guo, 2023). Previous research has confirmed the insight advantage effect in item memory (e.g., answers) (e.g., Chen et al., 2021; Kizilirmak et al., 2016, 2021; Shen et al., 2021, 2022). However, the essence of creative insight lies in establishing novel associations (Chen et al., 2021; Zhao et al., 2017). Thus, internal associations between elements are more important for insight events than individual items. In this study, we found that individuals performed better in associative recognition under high (vs. low) novelty conditions. This demonstrates the insight advantage effect in associative memory.

The question of why insight shows a memory advantage effect has been explored in recent research. One explanation suggests that novel associations, as a fundamental property of creative insight (Huang et al., 2018), are key to the memory advantage effect of insight (Chen et al., 2021; Zhao et al., 2024). Previous studies have indirectly or directly supported the role of novelty in item memory effects (Chen et al., 2021; Ding et al., 2023, 2024; Zhao et al., 2024). Chen et al. (2021) discovered that after one week, novel responses exhibited superior recognition performance compared to routine responses, a phenomenon associated with increased activation in the right hippocampus. Consistently, Zhao et al. (2024) found better memory performance was associated with longer fixation retentions for novel solutions. Ding et al. (2023, 2024) also found in the alternative use task that novelty ratings positively predicted recall and recognition performance, both of which were associated with activation of the novelty-related hippocampus.

This study directly reveals the promoting effect of novelty in associative memory within creative insight. Specifically, this study found an interaction effect between novelty and time: there was no significant difference between the high and low novelty groups in immediate testing after 3 min, but in delayed testing after 24 h, the high-novelty group had higher associative recognition rates than the low-novelty group. Additionally, we found consistent results that the effect of recognition time on hit rate, associative recognition rate, and the modified associative recognition rate, is all significant for low-novelty riddles but there are no significant differences between immediate testing and delayed testing for the high-novelty condition. This indicates that associative novelty promotes long-term memory and regulates forgetting depending on the duration of time: as time passes, associative novelty may enhance memory traces through consolidation of long-term memory, thereby slowing down forgetting. This promoting effect may be due to novel associations activating the hippocampus, enhancing consolidation of long-term memory, and thus slowing down forgetting. Previous research has shown that processing novelty in insight mainly activates the hippocampus (Chen et al., 2021; Milivojevic et al., 2015), and studies in the memory field have confirmed that associative memory mainly relies on hippocampal activity (Davachi, 2006; Diana et al., 2010). Consistent with this, Chen (2021) revealed that better recognition performance of novel (compared to routine) responses after one week is indeed related to activations in the right hippocampus.

Novelty plays different roles in associative memory and item memory. First, it is well known that the core of creative insight is building novel associations (Bowden et al., 2005; Milivojevic et al., 2015), which are processed with priority and stored in memory for later retrieval. Thus, novelty may be more beneficial for associative memory than item memory. Further studies can examine this. Second, the influence of novelty on associative memory and item memory may share common and different neuro-cognitive bases. On one hand, previous studies have indicated that the novelty of insight mainly involves the activities of the medial temporal lobe and the prefrontal cortex (Bowden et al., 2005; Chen et al., 2021; Huang et al., 2018; Milivojevic et al., 2015). On the other hand, previous studies have indicated that both associative memory and item memory require recollection (e.g., Yonelinas, 1997), which mainly relies on the activities of the hippocampus and prefrontal cortex (Davachi, 2006; Diana et al., 2010; Yonelinas, 2002). Item memory independently involves familiarity (e.g., Yonelinas, 1997), which mainly activates the perirhinal cortex (Davachi, 2006; Diana et al., 2010; Yonelinas, 2002). Although previous studies have primarily found that the hippocampus predicts better item memory (Chen et al., 2021; Ding et al., 2024), further studies can compare the memory differences of high- and low- novelty on associations and items.

This study expands on the novelty advantage effect of insight to the metamemory level. Previous research by Ludmer et al. (2011) showed that individuals exhibit a higher confidence in distinguishing correct from incorrect answers. Similarly, Shen et al. (2021) demonstrated that recognition confidence is greater for creative advertisements compared to conventional ones. The current study uncovers that individuals display increased confidence in the associative memory of "Mian"-"Midi" pairs under high-novelty conditions as opposed to low-novelty conditions. This suggests that insight novelty not only boosts item and associative recognition but also enhances confidence judgments in object (associative) memory. The impact of novelty and insight on metamemory may be mediated by the executive control system. Studies have indicated that novelty processing activates brain regions associated with cognitive monitoring, such as the dorsolateral prefrontal cortex and anterior cingulate cortex (Huang et al., 2015, 2018), which are also crucial for metamemory assessments (Fleming & Dolan, 2012). Future research should integrate brain imaging techniques to delve deeper into this phenomenon.

This study has both theoretical and practical implications. In theory, this study not only extends the promotion of induced insight on memory from item memory to associative memory and its metamemory, but also reveals the critical role of the novelty components underlying the promotion effect. Moreover, this study has critical implications for educational practice: compared to conventional or low-novelty learning, learning through insight by building high-novelty associations can enhance memory performance, thus providing an important pathway for efficient learning.

However, there are still some limitations. First, the recognition of associations between "Mimian"-"Midi" pairs may be affected by matching strategies during the testing phase, where participants base their old/new judgments on the match between the pairs rather than on familiarity or recollection. This could complicate the assessment of association rates, typically determined by comparing old pairs with rearranged ones, although this issue is consistent across both high-novelty and low-novelty scenarios. To address this, we introduced an equal number of new "Mimian"-"Midi" pairs, comprising both matched and non-matching pairs, for the old-new recognition judgment. The notable difference in false alarm rates between new matching and non-matching logographs indicates the impact of the matching strategy on memory performance. Furthermore, the incorporation of new logograph pairs enabled us to employ a modified formula for calculating the associative recognition rate, thus mitigating the influence of the matching strategy. Importantly, consistent results were obtained for memory performance across various metrics, including hit rates, association rates, and modified association rates. Despite our meticulous experimental design and measurements to minimize and detect this potential influence, caution is warranted when interpreting the data, and we recommend exploring enhanced methods and measurements in future research endeavors.

Second, another limitation is that the reliability score of novelty assessment is relatively low. Specifically, although we have distinguished low-novelty and high-novelty riddles based on the assessment by 20 raters, the value of the intraclass correlation coefficients (ICC = 0.56) falls in the range of moderate reliability according to previous guidelines [the ICC value is moderate in the range of 0.5–0.75 according to Koo & Li, 2016]. Considering that the reliability of creativity assessment may be influenced by the number of judges (Cseh & Jeffries, 2019; Long & Wang, 2022), we had ten more judges (for a total of 30 judges) rate the novelty (as well as comprehensibility and appropriateness) of the riddles. The post hoc complementary analysis showed moderate reliability for novelty as well (the ICC value is 0.69; 0.90 for comprehensibility and 0.88 for appropriateness). In fact, creativity, particularly its originality/novelty component, is variable and sensitive to complex factors, which makes it not easy to assess and often results in poor reliability (Benedek et al., 2016; Cseh & Jeffries, 2019; Forthmann et al., 2020; Long & Wang, 2022; Plucker et al., 2011; Reiter-Palmon et al., 2019). For example, Long and Wang (2022) reviewed 50 studies that involved the assessment of creativity with ICC's inter-judge reliability. They found that the ICC values ranged from 0.04 to 0.98, of which ICC values were lower than 0.30 in 4 % of studies, 0.31 to 0.50 in 4 % of studies, and 0.51 to 0.70 in 26 % of studies. Similarly, Plucker et al. (2011) compared three subjective rating methods in assessing originality/novelty, and they found low to moderate reliability as well (the ICC values are 0.62, 0.53, and 0.54, respectively). In short, further studies are suggested to examine and replicate the conclusions of this study.

Third, the research is predominantly grounded in experimental evidence using word logograph materials. It is essential to explore whether these results can be extrapolated to other insight materials and confirmed in educational environments beyond controlled settings. Fourth, while the study delves into the significance of novelty in the memory enhancement linked to induced insight, it is crucial to acknowledge that insight can also manifest spontaneously (Rothmaler et al., 2017). Previous studies have found a memory advantage effect of spontaneous insight in item memory (Chen et al., 2023; Cui et al., 2021; Danek et al., 2013; Danek & Wiley, 2020; Du et al., 2022). Therefore, future research can examine the role of novelty or appropriateness in the memory advantage effect of spontaneous insight.

CRediT authorship contribution statement

Zhonglu Zhang: Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Data curation, Conceptualization. **Yuye Si:** Writing – review & editing, Writing – original draft, Visualization, Validation, Data curation. **Lan Wang:** Software, Resources, Methodology, Conceptualization. **Kaiyun He:** Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

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

Acknowledgements

This work was supported by the MOE (Ministry of Education in China) Project on Humanities and Social Sciences (Grant No. 21YJC190025) and the National Natural Science Foundation of China (Grant No. 32271118). The authors declared that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2025.104901.

Data availability

Data will be made available on request.

References

Auble, P. M., Franks, J. J., & Soraci, S. A. (1979). Effort toward comprehension: Elaboration or "aha!"? *Memory & Cognition, 7*(6), 426–434. https://doi.org/10.3758/bf03198259

Bacso, S. A., & Marmurek, H. H. (2016). Testing effects of free recall on organization in whole/part and part/whole transfer. *Acta Psychologica, 171*, 99–109. https://doi.org/10.1016/j.actpsy.2016.10.001

Bader, R., Mecklinger, A., Hoppstädter, M., & Meyer, P. (2010). Recognition memory for one-trial-unitized word pairs: Evidence from event-related potentials. *NeuroImage, 50*(2), 772–781. https://doi.org/10.1016/j.neuroimage.2009.12.100

Benedek, M., Nordtvedt, N., Jauk, E., Koschmieder, C., Pretsch, J., Krammer, G., & Neubauer, A. C. (2016). Assessment of creativity evaluation skills: A psychometric investigation in prospective teachers. *Thinking Skills and Creativity, 21*, 75–84. https://doi.org/10.1016/j.tsc.2016.05.007

Bowden, E. M., & Jung-Beeman, M. (2007). Methods for investigating the neural components of insight. *Methods, 42*(1), 87–99. https://doi.org/10.1016/j.ymeth.2006.11.007

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences, 9*(7), 322–328. https://doi.org/10.1016/j.tics.2005.05.012

Buchler, N. G., Light, L. L., & Reder, L. M. (2008). Memory for items and associations: Distinct representations and processes in associative recognition. *Journal of Memory and Language, 59*(2), 183–199. https://doi.org/10.1016/j.jml.2008.04.001

Campbell, J. I., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods, 44*(4), 1255–1265. https://doi.org/10.3758/s13428-012-0186-0

Chen, J., Zhang, K., Du, X., Pan, J., & Luo, J. (2023). The neural mechanisms of the effect of spontaneous insight on re-solution: An ERP study. *Journal of Intelligence, 11*(1), 10. https://doi.org/10.3390/jintelligence11010010

Chen, S., Liang, Z., Li, X. L., Chen, Y. R., Zhao, Q. B., Yu, Q. L., Li, S. Q., Zhao, Z. J., & Liu, L. Z. (2021). The role of novel semantic association in the promoting effect of insight on memory. *Acta Psychologica Sinica, 53*(8), 837–846. https://doi.org/10.3724/sp.j.1041.2021.00837

Chrabaszcz, A., Gebremedhen, N. I., Alvarez, T. A., Durisko, C., & Fiez, J. A. (2023). Orthographic learning in adults through overt and covert reading. *Acta Psychologica, 241*, Article 104061. https://doi.org/10.1016/j.actpsy.2023.104061

Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009). Neural correlates of metamemory: A comparison of feeling-of-knowing and retrospective confidence judgments. *Journal of Cognitive Neuroscience, 21*(9), 1751–1765. https://doi.org/10.1162/jocn.2009.21123

Crouse, M. D. (2024). Learning to direct attention: Consequences for procedural task training programs. *Acta Psychologica, 250*, Article 104502. https://doi.org/10.1016/j.actpsy.2024.104502

Cseh, G. M., & Jeffries, K. K. (2019). A scattered CAT: A critical evaluation of the consensual assessment technique for creativity research. *Psychology of Aesthetics, Creativity, and the Arts, 13*(2), 159–166. https://doi.org/10.1037/aca0000220

Cui, C., Zhang, K., X., M., Sun, X., & Luo, J. (2021). Event-related potentials support the mnemonic effect of spontaneous insight solution. *Psychological Research, 85*(7), 2518–2529. https://doi.org/10.1007/s00426-020-01421-1

Danek, A. H., Fraps, T., von Müller, A., Grothe, B., & Öllinger, M. (2013). Aha! Experiences leave a mark: Facilitated recall of insight solutions. *Psychological Research, 77*(5), 659–669. https://doi.org/10.1007/s00426-012-0454-8

Danek, A. H., & Wiley, J. (2020). What causes the insight memory advantage? *Cognition, 205*, Article 104411. https://doi.org/10.1016/j.cognition.2020.104411

Danek, A. H., & Wiley, J. (2024). The insight memory advantage. In C. Salvi, J. Wiley, & S. M. Smith (Eds.), *The emergence of insight* (pp. 199–220). Cambridge: Cambridge University Press https://doi.org/10.1017/9781009244244.010.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology, 16*(6), 693–700. https://doi.org/10.1016/j.conb.2006.10.012

Diana, R. A., Yonelinas, R. P., & Ranganath, R. (2010). Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *Journal of Cognitive Neuroscience, 22*(8), 1808–1818. https://doi.org/10.1162/jocn.2009.21335

Ding, K., Chen, Q., Yang, W., et al. (2023). Recognizing ideas generated in a creative thinking task: Effect of the subjective novelty. *Current Psychology, 42*, 529–541. https://doi.org/10.1007/s12144-020-01342-7

Ding, K., He, R., Wang, X., Chen, Q., & Kenett, Y. N. (2024). Recognizing ideas generated in a creative task: The roles of the hippocampus and medial prefrontal cortex in facilitating self-generated learning. *Cerebral Cortex (New York, N.Y. : 1991), 34*(5), Article bhae219. https://doi.org/10.1093/cercor/bhae219

Dougherty, M. R., Scheck, P., Nelson, T. O., & Narens, L. (2005). Using the past to predict the future. *Memory & Cognition, 33*(6), 1096–1115. https://doi.org/10.1126/science.1214828

Du, X., Cui, C., Hu, Z., Zhang, K., & Song, Y. (2022). The mnemonic effects of insight on false memory in the DRM paradigm. *Psychological Research, 86*(2). https://doi.org/10.1007/s00426-021-01513-6, 558–57.

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 367*(1594), 1338–1349. https://doi.org/10.1007/978-3-642-45190-4_11

Forester, G., & Kamp, S. M. (2023). Pre-associative item encoding influences associative memory: Behavioral and ERP evidence. *Cognitive, Affective, & Behavioral Neuroscience, 23*(4), 1059–1075. https://doi.org/10.3758/s13415-023-01102-7

Forthmann, B., Paek, S. H., Dumas, D., Barbot, B., & Holling, H. (2020). Scrutinizing the basis of originality in divergent thinking tests: On the measurement precision of response propensity estimates. *British Journal of Educational Psychology, 90*(3), 683–699. https://doi.org/10.1111/bjep.12325

Hu, X., Yang, C., & Luo, L. (2022). Retrospective confidence rating about memory performance is affected by both retrieval fluency and non-decision time. *Metacognition and Learning, 17*, 651–681. https://doi.org/10.1007/s11409-022-09303-0

Huang, F., Tang, S., Sun, P., & Luo, J. (2018). Neural correlates of novelty and appropriateness processing in externally induced constraint relaxation. *Neuroimage, 172*, 381–389. https://doi.org/10.1016/j.neuroimage.2018.01.070

Huang, F. R., Fan, J., & Luo, J. (2015). The neural basis of novelty and appropriateness in processing of creative chunk decomposition. *Neuroimage, 113*, 122–132. https://doi.org/10.1016/j.neuroimage.2015.03.030

Kizilirmak, J., Fischer, L., Krause, J. J., Soch, J., Richter, A., & Schott, B. (2021). Learning by insight-like sudden comprehension as a potential strategy to improve memory encoding in older adults. *Frontiers in Aging Neuroscience, 13*, 661346. https://doi.org/10.3389/fnagi.2021.661346

Kizilirmak, J. M., Thuerich, H., Folta-Schoofs, K., Schott, B. H., & Richardson-Klavehn, A. (2016). Neural correlates of learning from induced insight: A case for reward-based episodic encoding. *Frontiers in Psychology, 7*, 1–16. https://doi.org/10.3389/fpsyg.2016.01693

Köhler, W. (1925). The mentality of apes. *Nature, 116*, 351–352. https://doi.org/10.1038/116351a0

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*(3), 490–517. https://doi.org/10.1037/0033-295x.103.3.490

Krieglstein, F., Jansen, S., Meusel, F., Scheller, N., Schmitz, M., Wesenberg, L., & Rey, G. D. (2024). Exploring fonts as retrieval cues in text-based learning. *Acta Psychologica, 251*, Article 104606. https://doi.org/10.1016/j.actpsy.2024.104606

Lai, Y. Q., Yang, Q., & Huang, B. Z. (2019). The promoting effect of insight on memory. *Advances in Psychological Science, 27*(12), 2034–2042. https://doi.org/10.3724/sp.j.1042.2019.02034

Li, D., Song, D., & Wang, T. (2020). Concreteness and imageability and their influences on Chinese two-character word recognition. *Reading and Writing, 33*, 1443–1476. https://doi.org/10.1007/s11145-020-10016-6

Long, H., & Wang, J. (2022). Dissecting reliability and validity evidence of subjective creativity assessment: A literature review. *Educational Psychology Review, 34*, 1399–1443. https://doi.org/10.1007/s10648-022-09679-0

Ludmer, R., Dudai, Y., & Rubin, N. (2011). Uncovering camouflage: Amygdala activation predicts long-term memory of induced perceptual insight. *Neuron, 69*(5), 1002–1014. https://doi.org/10.1016/j.neuron.2011.02.013

Mai, X. Q., Luo, J., Wu, J. H., & Luo, Y. J. (2004). "Aha!" effects in a guessing riddle task: An event-related potential study. *Human Brain Mapping, 22*(4), 261–270. https://doi.org/10.1002/hbm.20030

Milivojevic, B., Vicente-Grabovetsky, A., & Doeller, C. F. (2015). Insight reconfigures hippocampal-prefrontal memories. *Current Biology : CB, 25*(7), 821–830. https://doi.org/10.1016/j.cub.2015.01.033

Mohanty, P. P., & Naveh–Benjamin, M. (2018). Mitigating the adverse effects of response deadline on recognition memory: Differential effects of semantic memory support on item and associative memory. *Journal of Memory and Language, 102*, 182–194. https://doi.org/10.1016/j.jml.2018.05.010

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1170–1187. https://doi.org/10.1037/0278-7393.26.5.1170

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation, 26*, 125–173. https://doi.org/10.1016/s0079-7421(08)60053-5

Plucker, J. A., Qian, M., & Wang, S. (2011). Is originality in the eye of the beholder? Comparison of scoring techniques in the assessment of divergent thinking. *The Journal of Creative Behavior, 45*, 1–22. https://doi.org/10.1002/j.2162-6057.2011.tb01081.x

Qiu, J., Li, H., Luo, Y., Chen, A., Zhang, F., Zhang, J., Yang, J., & Zhang, Q. (2006). Brain mechanism of cognitive conflict in a guessing Chinese logograph task. *NeuroReport, 17*(6), 679–682. https://doi.org/10.1097/00001756-200604240-00025

Reiter-Palmon, R., Forthmann, B., & Barbot, B. (2019). Scoring divergent thinking tests: A review and systematic framework. *Psychology of Aesthetics, Creativity, and the Arts, 13*(2), 144–152. https://doi.org/10.1037/aca0000227

Robey, A. M., Castillo, C., Ha, J., Kerlow, M., Tesfa, N., & Dougherty, M. (2022). Generalizing the effect of type of metacognitive judgment on restudy decisions. *Metacognition and Learning, 17*, 73–85. https://doi.org/10.1007/s11409-021-09274-8

Robey, A. M., Dougherty, M. R., & Buttaccio, D. R. (2017). Making retrospective confidence judgments improves learners' ability to decide what not to study. *Psychological Science, 28*(11), 1683–1693. https://doi.org/10.1177/0956797617718800

Rothmaler, K., Nigbur, R., & Ivanova, G. (2017). New insights into insight: Neurophysiological correlates of the difference between the intrinsic "aha" and the extrinsic "oh yes" moment. *Neuropsychologia, 95*, 204–214. https://doi.org/10.1016/j.neuropsychologia.2016.12.017

Shen, W., Bai, H., Ball, L. J., Yuan, Y., & Wang, M. (2021). What makes creative advertisements memorable? The role of insight. *Psychological Research, 85*(7), 2538–2552. https://doi.org/10.1007/s00426-020-01439-5

Shen, W., Bai, H., Yuan, Y., Ball, L. J., & Lu, F. (2022). Quantifying the roles of conscious and unconscious processing in insight-related memory effectiveness within standard and creative advertising. *Psychological Research, 86*(5), 1410–1425. https://doi.org/10.1007/s00426-021-01572-9

Wills, T. W., Soraci, S. A., Chechile, R. A., & Taylor, H. A. (2000). "Aha" effects in the generation of pictures. *Memory & Cognition, 28*(6), 939–948. https://doi.org/10.3758/bf03209341

Wixted, J. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55*, 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Wu, X., Liu, Y., & Luo, J. (2018). The mnemonic effects of novelty and appropriateness in creative chunk decomposition tasks. *Frontiers in Psychology, 9*, 673. https://doi.org/10.3389/fpsyg.2018.00673

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition, 25*(6), 747–763. https://doi.org/10.3758/bf03211318

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus, 20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864

Zhang, J., & Liu, C. (2021). Brain activity in the process of verse recognition during the appreciation of Chinese classical poetry: An ERPs study. *Psychological Research (in Chinese), 14*(04), 297–304.

Zhang, Z., Deng, J., Zhang, M., Lu, Z., Zhou, Y., Lin, Y., & Xing, Q. (2023). The involvement of conflict monitoring and rewards processing in the appropriateness evaluation of creativity: An event-related potential-based analysis. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000659. Advance online publication.

Zhao, C. Y., & Guo, C. Y. (2023). Relationship between associative encoding and item encoding in the multiple-component character unitization and compound word unitization. *Acta Psychologica Sinica, 55*(4), 513–528. https://doi.org/10.3724/sp.j.1041.2023.00513

Zhao, Q., Guo, F., Chen, X., Chen, Y., Liang, Z., Yu, Q., & Zhou, Z. (2024). The advantage of novel solutions on subsequent memory in insight problems. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000662 *(Online publication)*.

Zhao, Q. B., Wei, L. L., Li, Y., Zhou, Z. J., Zhao, L. L., & Tang, L. (2017). Right hemispheric dominance in forming novel semantic associations. *Acta Psychologica Sinica, 49*(11), 1370–1382. https://doi.org/10.3724/sp.j.1041.2017.01370

Zhou, Y. (2022). *A study on the advantageous effects of item memory and associative memory for evoked insight*. Master's thesis (in Chinese). Guangzhou University. https://doi.org/10.27040/d.cnki.ggzdu.2022.001229.