A Dual-Consciousness Mechanism by Which False Memories Influence Decision Preferences*

Lu Feifei, Yan Tong, Zhou Chu, Wang Jianqin

2026-08-26

ChinaRxiv

Abstract

The constructive nature of memory makes it prone to biases during encoding and retrieval, resulting in false memories; however, whether false memories directly affect value-based decision-making and the underlying conscious mechanisms remain unclear. By integrating the Deese-Roediger-McDermott (DRM) paradigm of false memory with a reward-learning task (Experiment 1), this study systematically examined the influence of false memory on value-based decision-making, and introduced a divided-attention task (Experiment 2) and a process dissociation procedure (Experiment 3) to explore the level of consciousness at which this influence operates. The results showed that: (1) the current paradigm reliably induced false reward memories for critical lures, and false reward memories significantly influenced subsequent value-based decision-making; (2) the divided-attention task significantly weakened value-based decision-making guided by true memory, but had a smaller effect on value-based decision-making guided by false memory, suggesting that unconscious components are involved when false memory guides value-based decision-making; (3) based on the independence assumption of conscious and unconscious components in the process dissociation model, operational estimates of the conscious and unconscious components were obtained separately, and both types of conscious components positively predicted decision preferences for critical lures, with the unconscious component accounting for a larger amount of independent explanatory variance. This study indicates that false memory and true memory rely on different evidential bases when guiding value-based decision-making, and that dual conscious mechanisms are involved when false memory guides value-based decision-making, providing new evidence for understanding how memory construction processes influence value-based decision-making.

Full Text

A Dual-Consciousness Mechanism by Which False Memories Influence Decision Preferences*

Lu Feifei Yan Tong Zhou Chu† Wang Jianqin†

Department of Psychology, Fudan University, Shanghai 200433

Abstract

The constructive nature of memory makes it prone to biases during encoding and retrieval, thereby producing false memories. However, whether false memories directly affect value-based decisions, and the conscious mechanisms through which they operate, remain unclear. By integrating the false-memory Deese-Roediger-McDermott (DRM) paradigm with a reward-learning task (Experiment 1), the present study systematically examined the influence of false memories on value-based decisions, and introduced a divided-attention task (Experiment 2) and a process-dissociation procedure (Experiment 3) to explore the level of consciousness at which this effect occurs.

The results showed that: (1) the present paradigm can stably induce false reward memories for critical lures, and false reward memories significantly influence subsequent value-based decisions; (2) the divided-attention task significantly weakens value-based decisions guided by correct memories, whereas it has a smaller effect on value-based decisions guided by false memories, suggesting that value-based decisions guided by false memories contain an unconscious component; (3) based on the assumption of independence between conscious and unconscious processes in the process-dissociation model, operational estimates of conscious and unconscious components were obtained separately, and both types of conscious components were found to positively predict decision preferences for critical lures, with the unconscious component showing greater independent explanatory power. The study indicates that false memories and true memories guiding value-based decisions have different evidential bases, and that a dual-consciousness mechanism exists when false memories guide value-based decisions, providing new evidence for understanding how the constructive process of memory affects value-based decision-making.

Keywords: false memory; decision preference; reward learning; DRM paradigm; process-dissociation procedure

---

Received: 2025-10-31

* Supported by the Ministry of Education Humanities and Social Sciences Research Projects (20YJA190012, 24YJA190016).

† Zhou Chu and Wang Jianqin are co-corresponding authors.

Corresponding authors: Zhou Chu, Email: zhouchu@fudan.edu.cn; Wang Jianqin, Email: wjq@fudan.edu.cn

1 Introduction

Everyday decisions are often influenced by episodic memory. For example, people may tend to repurchase a product because they remember that a certain brand of chocolate had a distinctive flavor. However, memory is constructive: its retrieval process is not always precise and error-free, and systematic biases may also give rise to false memories (Deese, 1959; Roediger & McDermott, 1995). This constructive feature means that, when making decisions, individuals are not limited to original single experiences. Similar to true memories, false memories can also guide value-based decisions (Wang et al., 2024). Yet whether the influence of false memory on value-based decisions depends more on conscious or unconscious mechanisms remains to be explored. A deeper understanding of how the conscious mechanisms of false memory affect value-based decisions is of great significance for revealing the constructive nature of memory and the intrinsic mechanisms by which it influences value-based decision-making.

Studies have shown that true memories can effectively guide value-based decisions (Murty et al., 2016; Wimmer & Shohamy, 2012). In value-based decisions, decision makers associate subjective value with options based on the expected reward of each option and make decisions after comparing the values of the options (Weilbächer & Gluth, 2017). Murty et al. (2016) demonstrated that associative memory for "item–value" supports value-based decisions: after reward-association learning, participants were more inclined to choose lotteries that had previously been associated with higher rewards rather than lotteries associated with lower rewards. Notably, adaptive decision-making emerges only when participants can recall the value of the lottery. This mechanism allows value information to spread within the memory system, providing broader value cues for subsequent decisions.

However, memory is not always accurate and error-free. Its constructive nature can lead to biases during encoding and retrieval, thereby producing false memories. False memory refers to the phenomenon in which an individual mistakenly believes that an event that was never actually experienced has occurred, or in which memory for an event is inconsistent with the actual situation (Deese, 1959; Roediger & McDermott, 1995). Individuals' decision-making processes may also depend on biased false memories. Wang et al. (2024) combined the classic DRM paradigm of false memory (Deese/Roediger-McDermott Paradigm) with a reward-learning task, having participants learn that some associated pictures brought rewards (e.g., "butter," "toast," "sandwich"), whereas other associated pictures had no reward ("jeep," "wheel," "fire truck"). The results showed that participants formed false reward memories for critical lures that had not appeared (i.e., that "bread" had brought a reward), leading them to prefer critical-lure pictures in subsequent decision tasks. Moreover, the more false reward memories there were, the stronger the decision preference for the critical lures. These findings indicate that, like true memory, false memory can influence individuals' value-based decisions about stimuli they have never experienced in new situations. Based on spreading activation theory (Collins & Loftus, 1975; Roediger et al., 2001), the constructive association hypothesis proposed by Wang et al. (2024) suggests that, because critical lures and reward values are synchronously activated during encoding, the memory system can automatically construct erroneous associations between critical lures and value, thereby forming false reward memories. This study found that false reward memory can significantly predict decision preferences for critical lures, revealing for the first time the influence of false memory on value-based decisions; however, the intrinsic mechanisms by which false reward memory affects decision preferences remain unclear.

A key question is whether the influence of memory on value-based decisions occurs through conscious or unconscious components of memory. Schneider et al. (2021) pointed out that episodic memories can be formed and retrieved under either conscious or unconscious encoding conditions, and that episodic-memory retrieval networks (such as the hippocampus) can be activated during encoding and retrieval under both conscious conditions. On the one hand, some studies have found that decisions are affected only when participants retrieve memories in a conscious state (Murty et al., 2016). People may strategically attend to valuable information (Knowlton & Castel, 2022) and process or encode it more deeply semantically (Cohen et al., 2017); thus, participants are more likely to form stronger memories for high-value items, thereby influencing decisions. Neuroimaging evidence also shows that value information activates brain regions related to deep semantic processing and reward processing, thereby facilitating memory processing of valuable items (Cohen et al., 2014). On the other hand, memory may guide value-based decisions automatically, bypassing conscious retrieval. When making decisions, participants do not necessarily retrieve memories at the conscious level; the activation of the value of decision items may be a rapid, automatic process. During the processing of value information, the hippocampus may be directly influenced by midbrain dopaminergic signals, thereby enhancing memory representations of high-value information and being activated without the need for conscious encoding (Knowlton & Castel, 2022). For example, the co-activation of midbrain reward regions—especially the ventral tegmental area (VTA)—and the medial temporal lobe (such as the hippocampus) is associated with the formation of reward memory for "item-value," suggesting that reward-memory processing involves an automatic encoding pathway (Bowen et al., 2020). Therefore, the mechanism by which reward memory influences subsequent decisions may also be based on automatic unconscious processes. In the DRM paradigm, the generation of false memories may involve conscious or unconscious processes during encoding and retrieval (Zhou Chu et al., 2007; Mather et al., 1997; Roediger et al., 1998). If the influence of false memory on value-based decision-making originates from unconscious processes, then critical lures automatically activate value at the unconscious level and thereby influence decisions. If this influence originates from conscious processes, then individuals may further influence decision preferences only after retrieving, at the conscious level, an erroneous association of "critical lure-value." To explore the cognitive mechanism by which false memory guides value-based decisions, the present study manipulates the level of attention during the reward-learning phase as well as the conscious and unconscious components in the dissociative memory test phase, aiming to systematically investigate the conscious mechanism by which false memory influences decision preferences.

First, are there specific differences between the mechanisms by which false memory and true memory guide value-based decisions? According to activation–diffusion theory, activation of studied items during the encoding phase can spread through the semantic network to critical lures, thereby increasing the tendency to falsely remember critical lures (Collins & Loftus, 1975; Roediger et al., 2001). In a reward-learning context, the constructive-association hypothesis further suggests that when studied items and rewards are co-activated, this may lead to a constructive link between critical lures and rewards, enabling critical lures to automatically activate reward representations in subsequent decisions and thereby influence value-based decisions (Wang et al., 2024). Therefore, if the value cues of critical lures rely more on automatic semantic associative activation and constructive linkage, whereas studied items rely more on item-specific evidence provided by conscious processing, then the two may differ in their sensitivity to the reduction of attentional resources. The divided-attention task (Peterson & Peterson, 1959) has been shown to occupy attentional resources by increasing participants' cognitive load, and this task mainly affects general attentional resources (general attentional resources; Brown &

Brockmole, 2010; Yin et al., 2024). Introducing an attention-diversion task into the DRM paradigm can effectively manipulate individuals' conscious processing of studied items and critical lures (Dewhurst et al., 2007; Otgaar et al., 2012), while its influence on unconscious processing is relatively weak or even absent (Elliott & Brewer, 2019; Palmer et al., 2010; Turk et al., 2013). Therefore, if the internal mechanisms by which false memories and true memories guide value-based decision-making differ, then after the conscious component of memory is weakened by the use of an attention-diversion task, value-based decisions guided by false memories will show differences from those guided by true memories; conversely, if the internal mechanisms by which the two guide value-based decision-making are similar, then introducing an attention-diversion task will make their patterns of influence on value-based decision-making similar.

Furthermore, if there are specific differences between the internal mechanisms by which false memories and true memories guide value-based decision-making, what are the deeper reasons behind them? According to activation monitoring theory (AMT; Roediger et al., 2001), the activation process and the monitoring process jointly influence the formation of false memories. During the encoding of semantically associated words, activation spreads from the representations of semantically associated words to the representations of critical lures; critical lures are the result of the continuous accumulation of unconscious activation (McDermott & Watson, 2001; Roediger, 2001). When the semantic activation signal for a critical lure is insufficiently monitored, it is more likely to be mistakenly judged as a studied item. This unconscious activation and monitoring failure may be the fundamental reason for the specific differences when false memories guide value-based decision-making. The study by Zhou Chu et al. (2007) found that even when studied items were processed unconsciously, the false-memory effect for critical lures was still observed, indicating that false memories can arise through relatively automatic processes and do not require the involvement of a conscious generation process; of course, this does not mean that conscious processes play no role in false memory, but perhaps only that the level of consciousness required is relatively low.

Memory theories hold that both conscious and unconscious components exist simultaneously in memory. For example, source memory is a conscious, attention-regulated process involving the detailed reinstatement of the learning context; unconscious components, by contrast, are the result of rapid, automatic processing and are not influenced by conscious participation (Jacoby et al., 1993; Mandler, 1980). Recent evidence suggests that contextual memory exists in both conscious and unconscious forms (Schneider et al., 2021). To separate the relative contributions of conscious control and unconscious automatic processing to memory retrieval, Jacoby et al. (1993), building on the dual-process model, proposed the Process-dissociation procedure (PDP). PDP assumes that conscious and unconscious processes are mutually independent processing processes. By comparing participants' task performance in inclusion and exclusion tests, the relative contributions of the two can be separated. Given the similarities between false memories and true memories (Jou & Flores, 2013; Osth et al., 2026), their covariation (Zhou Chu et al., 2007), and features such as shared semantic or surface characteristics (Chang & Brainerd, 2021; Osth et al., 2026), as well as the operability of the PDP framework in false-memory research (Failes et al., 2020; Jacoby et al., 2005), the present study used the PDP to separately calculate the functional contributions of conscious processing and unconscious processing, estimate the conscious and unconscious components of false memory, and separately examine their effects on decision preferences, so as to further answer the conscious mechanisms by which false memory influences decision preferences.

In sum, concerning the key theoretical question of "how false reward memory affects value-based decision-making," existing research has mostly remained at its···

whether the phenomenal level of decision-making is affected, while there remains a lack of exploration of the conscious mechanisms on which it depends. Based on the constructive-association hypothesis and activation-monitoring theory, this study combines the attention-dispersion task with a process-dissociation procedure to directly separate the relative contributions of conscious and unconscious components when false-memory-guided value-based decisions are made, thereby clarifying for the first time whether false reward memories guide value-based decisions through conscious or unconscious components. The study by Wang et al. (2024) first found that false memories can guide value-based decisions; however, because the participants were all Western samples, this conclusion still requires replication. Therefore, Experiment 1 refers to the experimental paradigm of Wang et al. (2024), combining the DRM paradigm with a reward-learning task to examine the influence of false reward memories on decision preferences. On the basis of Experiment 1, Experiment 2 introduces an attention-dispersion task to manipulate the level of false memory and explores whether the influence of false memories on decision preferences depends on conscious processing. Because attention dispersion occupies participants' cognitive resources (Peterson & Peterson, 1959) and affects participants' extraction of critical lures at the conscious level (Dewhurst et al., 2007; Otgaar et al., 2012), Experiment 2 hypothesizes that attention dispersion can weaken subsequent decision preferences for critical lures under the rewarded condition; that is, the influence of false memories on decision preferences involves processing at the conscious level. Experiment 3 innovatively introduces a process-dissociation procedure to separate the conscious and unconscious components of false memories, and further investigates the predictive effects of the two conscious components on decision preferences, in order to examine in depth the cognitive mechanism by which false memories influence value-based decisions. According to activation-monitoring theory, critical lures may be the result of accumulated continuous unconscious activation (McDermott & Watson, 2001; Roediger, 2001); therefore, Experiment 3 hypothesizes that unconscious components in false reward memories can also predict decision preferences.

2 Experiment 1: The Influence of False Memories on Value-Based Decision Preferences

2.1 Method

2.1.1 Participants In the study by Wang et al. (2024), participants' decision preference for critical lures under rewarded conditions was significantly greater than the chance level (50%), with a relatively large effect size ($d = 0.99$). Therefore, Experiment 1 used G*Power 3.0 (Faul et al., 2007), conservatively setting a medium effect size of $d = 0.5$, $\alpha = 0.05$, and statistical power of power $= 0.80$, which indicated that 34 participants needed to be recruited. In the end, a total of 37 participants were recruited ($M = 20.38$, $SD = 2.48$, 15 males). All participants had normal or corrected-to-normal vision, volunteered to participate, and signed an informed consent form before the formal experiment began. They received monetary compensation upon completing the experiment.

2.1.2 Experimental Materials The materials included 10 sets of DRM picture lists and 30 unrelated neutral pictures. Each DRM picture list contained 8 studied items (e.g., "butter," "toast," "sandwich"), 1 critical lure (e.g., "bread"), and 1 list-related item (e.g., "milk"). Among them, the list-related items and critical lures were presented only during the memory test phase, and belonged to the studied items ···

For pictures of the same category, the listed related items had a relatively low association with the DRM list pictures, whereas the critical lures had a relatively high association with the DRM list pictures. The 10 DRM picture lists were randomly and evenly assigned to the rewarded and unrewarded conditions to balance the associative strength of the picture lists. The rewarded and unrewarded picture lists were also balanced across different participants. Taking the theme picture "bread" as an example, half of the participants learned that bread was associated with reward, whereas the other half learned that bread was associated with no reward. The above picture materials were all taken from Wang et al. (2018, 2024), the corresponding word lists were from Stadler et al. (1999), and the associative strength between the critical lures and the studied items ranged from 0.11 to 0.35 (Roediger et al., 2001). There were two groups of practice materials, each containing 2 studied items. The pictures in the practice materials were unrelated to those in the formal materials and were used for practice preparation in the memory-test phase.

2.1.3 Experimental Design and Procedure

Experiment 1 used a one-factor within-participants experimental design. The independent variable was the reward condition of the DRM list, divided into rewarded and unrewarded. The dependent variables included: (1) the recognition rate in the memory-test phase: correct recognition and false recognition. The former refers to participants correctly judging studied items as old pictures that had appeared, and the latter refers to participants incorrectly judging critical lures or list-related items that had not appeared as old pictures that had appeared; (2) the reward memory rate in the memory-test phase: correct reward memory and false reward memory. The former refers to participants correctly judging that studied items were rewarded, and the latter refers to participants incorrectly judging that critical lures or list-related items were rewarded; (3) in the decision phase: decision preferences for different types of pictures; and (4) the correlation analysis between reward memory and decision preferences.

Following the study by Wang et al. (2024), the DRM paradigm (Roediger & McDermott, 1995) was combined with a reward-learning task (Doll et al., 2015; Wimmer & Shohamy, 2012). The experiment was divided into three phases: the reward-learning phase, the memory-test phase, and the decision phase (see Figure 1).

In the reward-learning phase, there were 5 rounds in total. In each round, the studied items from two DRM picture lists were presented in random order; one group was rewarded and the other group was unrewarded, and the two groups of pictures were randomly intermixed. The DRM lists were balanced between odd- and even-numbered participants and between rewarded and unrewarded conditions. Each time, one picture was presented on the screen, and participants were required to predict whether each picture was rewarded: press the "Y" key for reward and the "N" key for no reward. After participants pressed a key, the correct answer was presented: if a picture of a five-yuan RMB note appeared, it indicated that the picture was rewarded; if a gray rectangle appeared, it indicated that the picture was unrewarded. There was no time limit for the presentation of each picture. After the participant pressed a key, the next picture appeared, separated by a 1000 ms fixation point. Before the experiment, participants were told that whether or not a picture was rewarded was an attribute of the picture itself and was not affected by their choices. Participants could enter the formal experiment only after completing the practice phase. The reward-learning phase and the memory-test phase were separated by 5 minutes, during which participants completed a matching game as a filler task.

In the memory-test phase, the studied items from each DRM picture list (positions 1, 4, and 7), the critical lures, the list-related items, and 30 unrelated items were presented in random order. Participants needed to judge: (1) whether the

Figure 1. Flowchart of Experiment 1

Figure 1: Figure 1. Flowchart of Experiment 1

picture had appeared in the reward-learning phase—press "1" for an old picture that had appeared, and press "2" for a new picture; (2) whether the picture had been associated with reward in the reward-learning phase—press "1" for rewarded pictures, press "2" for unrewarded pictures, press "3" if it had not appeared, and press "4" if they did not remember.

decision stage. One picture was presented on each side of the screen, and the participant' s task was to choose the picture that was more likely to receive a reward (Wang et al., 2019; Wimmer & Shohamy, 2012). If the picture on the left was rewarded, they pressed the "F" key; if the picture on the right was rewarded, they pressed the "J" key. There was no time limit for the choice. Each pair of pictures was presented four times in a random order, with the number of presentations of the pictures on the left and right balanced. Participants needed to make decisions among three types of picture combinations: (1) learned items with reward versus no reward, (2) critical-lure items versus list-related items, and (3) learned items versus list-related items.

Figure 1. Flowchart of Experiment 1

Note: This flowchart is a schematic illustration. In Panel B, pictures with dashed borders are critical lures, pictures with solid borders are learned items, and pictures with dotted borders are unrelated items.

2.2 Results

The data were analyzed using RStudio (4.4.0), with the bruceR (Bao, 2021) and lme4 (Bates et al., 2015) packages. For the analysis-of-variance results, Mauchly' s test of sphericity was first used to examine whether the data met the sphericity assumption; when the sphericity assumption was violated, the Greenhouse-Geisser correction was applied to the p value. Effect sizes were reported using η_p^2 and Cohen' s d. Post hoc comparisons were corrected using Bonferroni correction to control the false-positive probability caused by multiple comparisons. The same data-analysis method was used in the subsequent experiments.

2.2.1 Reward-learning stage: validation of the effectiveness of reward learning

Based on participants' keypresses during the reward-learning stage, the probability of judging a picture as rewarded was calculated separately under the reward and no-reward conditions at different serial positions (1–8) of the DRM pictures (see Figure 2A). The results showed that the main effect of reward condition was significant, $F(1, 36) = 65.17, p < 0.001, \eta_p^2 = 0.64, 95\%$ CI $= [0.48, 0.75]$; the main effect of presentation position was not significant, $F < 1$; and the interaction between the two was significant, $F(4.43, 159.47) = 12.40, p < 0.001, \eta_p^2 = 0.26, 95\%\ \mathrm{CI} = [0.15, 0.34]$. The later the trial position, the better participants were able to correctly distinguish rewarded from unrewarded pictures, indicating that participants had learned the association between pictures and rewards during the reward-learning stage.

Figure 2. Results of Experiment 1. (A) Learning curves in the reward-learning phase: the reward-prediction accuracy judged by participants at different test positions under rewarded and unrewarded conditions. **(B)** Memory-test phase: correct recognition rates for learned items and false recognition rates for key lures and list-related items; correct reward-memory rates for learned items and false reward-memory rates for key lures and list-related items. **(C)** Decision preferences in the decision phase: decision preference for rewarded learned items, and decision preference for rewarded key lures under the rewarded condition. **(D)** Correlation analysis between reward memory and decision preference: the fitted lines in the scatterplots show that correct reward memory predicted the decision preference for rewarded learned items, and false reward memory predicted the decision preference for key lures under the rewarded condition. Note: Error bars represent confidence intervals (95% CI), *** $p < 0.001$, n.s. indicates no significant difference.

2.2.2 Memory-test phase: recognition and reward memory

(1) Correct recognition and false recognition

The recognition rates for different item types are shown in Table 1. A 2 (reward condition: rewarded, unrewarded) × 3 (item type: learned items, key lures, and list-related items) two-factor repeated-measures analysis of variance was used to further compare differences in recognition judgments for different item types under different reward conditions. The results showed that the main effect of reward and the two-factor interaction were not significant, $Fs < 1$; the main effect of item type was significant, $F(1.55, 55.63) = 252.85, p < 0.001, \eta_p^2 = 0.88, 95\%\mathrm{CI} = [0.83, 0.91]$. To understand participants' levels of correct recognition and false recognition, one-way analyses of variance were conducted separately for rewarded and unrewarded items, comparing the correct recognition rate for learned items with the false recognition rates for key lures and list-related items. The results are shown in Figure 2B-left, and the descriptive results are shown in Table 1.

Table 1. Recognition rates and reward-memory rates for different item types ($M \pm SD$)

Reward condition	Recognition rate: Learned items	Recognition rate: Key lures	Recognition rate: List-related items	Reward-memory rate: Learned items	Reward-memory rate: Key lures	Reward-memory rate: List-related items
Rewarded	0.91 ± 0.10	0.47 ± 0.31	0.10 ± 0.13	0.70 ± 0.23	0.35 ± 0.29	0.04 ± 0.08
Unrewarded	0.88 ± 0.15	0.48 ± 0.31	0.06 ± 0.12	0.67 ± 0.22	0.36 ± 0.28	0.03 ± 0.08

For the recognition rates of rewarded items (item type: learned items, key lures, and list-related items), a one-factor···

A one-way analysis of variance was conducted on the recognition rates for rewarded items (item type: studied items, critical lures, and list-related items). The results showed a significant main effect of item type, $F(1.45, 52.04) = 156.52, p < 0.001, \eta_p^2 = 0.81$, 95% CI = [0.74, 0.86]. Post hoc tests revealed that the correct recognition rate for rewarded studied items was significantly higher than the false recognition rate for list-related items ($M = 0.10, SD = 0.13$), $t(36) = 26.20, p_{bonf} < 0.001$, Cohen' s $d = 2.91$, 95% CI = [2.63, 3.18]; the false recognition rate for critical lures in the rewarded condition was significantly higher than the false recognition rate for list-related items, $t(36) = 8.00, p_{bonf} < 0.001$, Cohen' s $d = 1.33$, 95% CI = [0.91, 1.74]; in the rewarded condition, the correct recognition rate for studied items was significantly higher than the false recognition rate for critical lures, $t(36) = 7.77, p_{bonf} < 0.001$, Cohen' s $d = 1.58$, 95% CI = [1.07, 2.09], indicating that participants' level of correct recognition for studied items was higher than their level of false recognition for critical lures.

A one-way analysis of variance was conducted on the recognition rates for non-rewarded items (item type: studied items, critical lures, and list-related items). The results showed a significant main effect of item type, $F(1.58, 56.69) = 168.45, p < 0.001, \eta_p^2 = 0.82$, 95% CI = [0.76, 0.87]. Post hoc tests revealed that the correct recognition rate for non-rewarded studied items was significantly higher than the false recognition rate for list-related items, $t(36) = 24.76, p_{bonf} < 0.001$, Cohen' s $d = 3.02$, 95% CI = [2.71, 3.32]; the false recognition rate for critical lures in the non-rewarded condition was also significantly higher than the false recognition rate for list-related items, $t(36) = 9.30, p_{bonf} < 0.001$, Cohen' s $d = 1.53$, 95% CI = [1.12, 1.94]; in the non-rewarded condition, the correct recognition rate for studied items was significantly higher than the false recognition rate for critical lures, $t(36) = 7.52, p_{bonf} < 0.001$, Cohen' s $d = 1.49$, 95% CI = [0.99, 1.99], indicating that participants' memory effect for studied items was better than that for critical lures. The above results indicate that participants produced false memories for critical lures under both rewarded and non-rewarded conditions.

(2) Correct reward memory and false reward memory

The reward memory rates for different item types are shown in Table 1. A 2 (reward condition: rewarded, non-rewarded) × 3 (item type: studied items, critical lures, and list-related items) two-factor repeated-measures analysis of variance was used to further compare differences in source judgments among different item types under different reward conditions. The results showed that the main effect of reward and the two-factor interaction were not significant, $Fs < 1$; the main effect of item type was significant, $F(2, 72) = 175.48, p < 0.001, \eta_{\mathrm{p}}^2 = 0.83$, 95% CI = [0.77, 0.87]. To understand participants' levels of correct reward memory and false reward memory, one-way analyses of variance were conducted separately for rewarded and non-rewarded items to compare differences between the correct reward memory for studied items and the false reward memories for critical lures and list-related items. The results are shown in Figure 2B-right.

For the memory rates of rewarded items (item type: studied items, critical lures, and list-related items), a one-way analysis of variance was conducted. The results showed a significant main effect of item type, $F(2, 72) = 109.35, p < 0.001, \eta_{\mathrm{p}}^2 = 0.75$, 95% CI = [0.67, 0.81]. Post hoc tests revealed that the correct reward memory for rewarded studied items was significantly higher than the false reward memory for list-related items, $t(36) = 16.41, p_{bonf} < 0.001$, Cohen' s $d = 2.43$, 95% CI = [2.06, 2.80]; the false reward memory for critical lures was significantly higher than the false reward memory for list-related items, $t(36) = 6.84, p_{bonf} < 0.001$, Cohen' s $d = 1.12$, 95% CI = [0.71, 1.53], indicating that erroneous memory for reward experiences could, through the DRM paradigm and reward learning

induced by the process; the correct reward memory for studied items was significantly higher than the false reward memory for critical lures, $t(36) = 7.28, p_{bonf} < 0.001$, Cohen' s $d = 1.31$, 95% CI = $[0.86, 1.76]$, indicating that although participants produced false reward memories for critical lures, their proportion was lower than that of correct reward memories.

A one-way ANOVA was conducted on the memory rate for non-rewarded items (item type: studied items, critical lures, and list-related items). The results showed a significant main effect of item type, $F(2, 72) = 126.46, p < 0.001, \eta_p^2 = 0.78$, 95% CI = $[0.71, 0.83]$. Post hoc tests showed that correct non-reward memory for non-rewarded studied items was significantly higher than false non-reward memory for list-related items, $t(36) = 18.04, p_{bonf} < 0.001$, Cohen' s $d = 2.61$, 95% CI = $[2.25, 2.98]$; false non-reward memory for critical lures was significantly higher than false non-reward memory for list-related items, $t(36) = 7.47, p_{bonf} < 0.001$, Cohen' s $d = 1.33$, 95% CI = $[0.88, 1.77]$, indicating that false memory for non-rewarded experiences can also be induced through the DRM paradigm and reward-learning process; the correct memory rate for non-rewarded studied items was significantly higher than the false memory rate for critical lures, $t(36) = 7.62, p_{bonf} < 0.001$, Cohen' s $d = 1.29$, 95% CI = $[0.86, 1.71]$, indicating that participants' memory for studied items was better than that for critical lures. The above results indicate that participants produced false memories for critical lures under both rewarded and non-rewarded conditions.

2.2.3 Decision Stage: Decision Preference

During the decision process, if participants had no preference for either of the two pictures, then the selection rate for the pictures would not differ significantly from the 50% chance level. The results of one-sample t tests are shown in Figure 2C. In choices between rewarded studied items and non-rewarded studied items, participants' decision preference for rewarded studied items ($M = 0.82, SD = 0.23$) was significantly higher than the 50% chance level, $t(36) = 8.28, p < 0.001$, Cohen' s $d = 1.36$, 95% CI $= [1.03, 1.69]$, again verifying that, during the reward-learning stage, participants learned the association between pictures and rewards.

In choices between critical lures and list-related items, participants' decision preference for critical lures under the rewarded condition ($M = 0.80, SD = 0.17$) was significantly higher than the 50% chance level, $t(36) = 10.51, p < 0.001$, Cohen's $d = 1.73$, 95% CI $= [1.39, 2.06]$, indicating that participants exhibited a decision preference. The selection rate for critical lures under the non-rewarded condition ($M = 0.44, SD = 0.26$) did not differ significantly from the 50% chance level, $t(36) = 1.30, p = 0.20$.

In choices between studied items and list-related items, participants' decision preference for rewarded studied items ($M = 0.77, SD = 0.22$) was significantly higher than the 50% chance level, $t(36) = 7.31, p < 0.001$, Cohen' s $d = 1.20$, 95% CI $= [0.87, 1.54]$, indicating that participants learned which pictures were rewarded during the reward-learning stage and could distinguish them at the decision stage. The selection rate for non-rewarded studied items ($M = 0.41, SD = 0.28$) did not differ significantly from the 50% chance level, $t(36) = 1.86, p = 0.07$.

2.2.4 Predictive Analysis of the Effect of Reward Memory on Decision Preference

To further explain, at the individual level, the relationship between decision preference for pictures and reward memory, the correct reward memory for rewarded studied items was used as a predictor.

Correct rewarded memory and false rewarded memory for the critical lure were correlated separately with decision preferences for the pictures. The results showed that correct rewarded memory significantly predicted decision preferences for items that had been rewarded during learning ($r_{(35)} = 0.51, p = 0.001$, Figure 2D-left), indicating that the higher the rewarded-memory rate for a picture, the stronger the decision preference shown at the decision stage. False rewarded memory likewise significantly predicted decision preferences for critical lures under the reward condition ($r_{(35)} = 0.43, p = 0.008$, Figure 2D-right), indicating that erroneous associative memory for picture value can effectively guide value-based decisions.

2.3 Discussion

Experiment 1 demonstrated that the DRM paradigm and a reward-learning task can effectively induce false rewarded memory for critical lures, and that false rewarded memory can effectively predict individuals' decision preferences for critical lures. Under the reward condition, participants showed false recognition and false rewarded memory for critical lures. This indicates that participants not only falsely remembered that the critical lures had appeared, but also falsely remembered that they were associated with reward, and that this was not the result of categorical inference. Specifically, false rewarded memory led participants to choose critical lures more often during the decision stage, showing a significant preference for critical lures, whereas false unrewarded memory prompted participants to avoid choosing critical lures during the decision stage. In Experiment 1, there was no difference between participants' correct and false memory effects for rewarded and unrewarded pictures. This is not entirely consistent with previous findings that reward can enhance participants' memory effects (Shohamy & Adcock, 2010), perhaps because participants engaged in active prediction and learning, with sufficient cognitive resources to process and learn the items, thereby processing the materials under the two conditions to the same degree.

It is known that false rewarded memory can guide decisions about novel, unexperienced stimuli, whereas false memory may involve conscious or unconscious retrieval processes (Zhou Chu et al., 2007; Mather et al., 1997; Roediger et al., 1998). Thus, does false memory guide value-based decisions at the conscious level or the unconscious level? Answering this question will help further reveal the internal mechanism by which false memory influences decision preferences. Introducing a divided-attention task into the false-memory DRM paradigm can manipulate individuals' processing of learned items and critical lures at the conscious rather than unconscious level (Dewhurst et al., 2007; Otgaar et al., 2012). If decision preferences are affected by conscious components, then divided attention may weaken or even eliminate decision preferences; whereas if decisions are affected by unconscious processing, divided attention should have difficulty influencing decision preferences based on false memory. Therefore, Experiment 2 manipulated the level of participants' false memory through divided attention to explore how the conscious processing level of false memory affects decision preferences.

3 Experiment 2: The Influence of False Memory on Value-Based Decision Preferences Under Different Attention Conditions

3.1 Method

3.1.1 Participants

Using G*Power 3.0 (Faul et al., 2007), and with reference to Umanath et al. (2019), to manipulate the influence of attention conditions on false memory

of the study, setting a medium effect size of $d = 0.60$, $\alpha = 0.05$, and statistical power power $= 0.80$; a total of 72 participants needed to be recruited. In practice, 73 participants were recruited, including 36 in the divided-attention group ($M = 22.64$, $SD = 1.81$; 10 males) and 37 in the full-attention group ($M = 21.78$, $SD = 2.04$; 12 males). All participants had normal or corrected-to-normal vision, participated voluntarily, and signed an informed-consent form before the formal experiment began. They received monetary compensation after completing the experiment.

3.1.2 Experimental Materials

Consistent with Experiment 1, both the formal materials and the practice materials were taken from Wang et al. (2018, 2024), comprising 10 sets of DRM word lists and 30 unrelated neutral pictures.

3.1.3 Experimental Design and Procedure

A 2 (reward condition: reward, no reward) × 2 (attention condition: divided attention, full attention) mixed experimental design was adopted, in which reward condition was a within-subjects variable and attention condition was a between-subjects variable. The dependent variables were the same as in Experiment 1.

Experiment 2 was a fully supervised online experiment. The experimental program was written using PsychoPy and uploaded to the Naodao research platform (Naodao-NAODAO) to generate an online experiment link. Throughout the experiment, participants were required to share their screen via Tencent Meeting and turn on their camera and microphone, with their mobile phone set to silent mode. After the experimenter confirmed that the participant was in a quiet, undisturbed environment and could hear the instructions normally, the online experiment began. Experiment 2 was likewise divided into three phases. In the reward-learning phase, participants' attentional state was manipulated; the remaining phases were the same as in Experiment 1.

During the reward-learning phase, participants were randomly assigned to the full-attention group or the divided-attention group. For the full-attention group, the presentation of stimuli was basically the same as in Experiment 1, differing only in the feedback provided for rewarded and unrewarded items. In the full-attention group, feedback was presented on the screen for 2000 ms. Participants only needed to observe the pattern in which pictures were presented and did not need to make keypress responses. A five-yuan RMB note indicated that the picture was rewarded, whereas a gray rectangle indicated that the picture was not rewarded. While completing the above task, the divided-attention group also had to perform a divided-attention task (Peterson & Peterson, 1959; Ricker et al., 2016), that is, count backward aloud starting from 300, subtracting 3 each time, such as "300 minus 3 equals 297, 297 minus 3 equals 294···" . If the count reached 0 or a negative number, participants restarted backward counting from a random number between 200 and 500 provided by the experimenter, continuing until the reward-learning phase ended. There was a 5-minute interval between the reward-learning phase and the memory-test phase. During the interval, participants were required to complete a 5-minute figural reasoning task as a filler task. The task materials were mental-rotation test items involving plane geometry or solid geometry, and all geometric figures were unrelated to the experimental materials.

3.2 Results

Before the results analysis, participants' false-recognition levels for unrelated items under the divided-attention and full-attention conditions were first compared. The results showed that false recognition of unrelated items under the divided-attention condition (M = 0.18, SD = 0.18) was significantly higher than under the full-attention condition (M = 0.04, SD = 0.07), $t(46.53) = 4.35$, $p < 0.001$, Cohen' s $d = 1.02$, 95% CI = [0.55, 1.50]. This indicates that divided attention affects the criterion participants use in making memory judgments. To more accurately reflect participants' memory effects and reduce the influence of guessing on the results,

effects. Before conducting data analysis in the memory-test phase, taking unrelated items as the baseline, the recognition results and reward memory for studied items, critical lures, and list-related items were corrected. Following previous methods for calculating corrected false memories (Brainerd et al., 2008; Wang et al., 2019): corrected recognition rate = pre-correction recognition rate − recognition rate for unrelated items; corrected reward memory rate = pre-correction reward memory rate − reward memory rate for unrelated items. If the corrected result was negative, this indicated that the participant had almost formed no memory and was purely guessing; therefore, it was recorded as 0. The analyses reported below are all based on the corrected data.

3.2.1 Memory-test phase: recognition and reward memory

(1) Correct recognition and false recognition

According to the experimental aims, the effects of reward and attention dispersion on recognition rates for different item types were analyzed; descriptive results are shown in Table 2. With the correct recognition rate for studied items as the dependent variable, a 2 (reward condition: absent, present) × 2 (attention condition: full attention, divided attention) repeated-measures ANOVA revealed a significant main effect of reward condition, $F(1, 71) = 15.68$, $p < 0.001$, $\eta_p^2 = 0.18$, 95% CI = [0.07, 0.31]. The correct recognition rate for rewarded studied items ($M = 0.71$, $SD = 0.20$) was significantly higher than that for unrewarded studied items ($M = 0.65$, $SD = 0.22$), indicating that participants learned rewarded studied items better. The main effect of attention condition was significant, $F(1, 71) = 40.73$, $p < 0.001$, $\eta_p^2 = 0.37$, 95% CI = [0.22, 0.49]. The correct recognition rate for studied items in the full-attention group ($M = 0.83$, $SD = 0.28$) was significantly higher than that in the divided-attention group ($M = 0.53$, $SD = 0.29$), indicating that divided attention reduced the learning effect for old pictures. The interaction between the two was not significant, $F < 1$.

Table 2. Recognition rates and reward-memory rates for different item types under full-attention and divided-attention conditions ($M \pm SD$)

		Full attention	Full attention	Full attention	Divided attention	Divided attention	Divided attention
		Studied items	Critical lures	List-related items	Studied items	Critical lures	List-related items
Item recognition rate	Rewarded	0.86 ± 0.12	0.39 ± 0.24	0.10 ± 0.19	0.56 ± 0.26	0.36 ± 0.27	0.12 ± 0.13
Item recognition rate	Unrewarded	0.80 ± 0.16	0.49 ± 0.27	0.08 ± 0.17	0.50 ± 0.28	0.46 ± 0.26	0.05 ± 0.12
Reward memory rate	Rewarded	0.77 ± 0.26	0.38 ± 0.26	0.09 ± 0.19	0.48 ± 0.31	0.28 ± 0.29	0.12 ± 0.12
Reward memory rate	Unrewarded	0.71 ± 0.22	0.42 ± 0.28	0.08 ± 0.19	0.37 ± 0.29	0.30 ± 0.27	0.08 ± 0.13

With the false recognition rate for critical lures as the dependent variable, a two-factor repeated-measures ANOVA with reward condition and attention condition found a significant main effect of reward condition, $F(1, 71) = 14.33$,

$p < 0.001$, $\eta_p^2 = 0.17$, 95% CI = [0.06, 0.30]. The false recognition rate for critical lures in the rewarded condition ($M = 0.38$, $SD = 0.25$) was significantly lower than that in the unrewarded condition ($M = 0.48$, $SD = 0.26$), indicating that participants were more likely to falsely recognize critical lures in the unrewarded condition. The other main effect and the interaction were not significant, $Fs < 1$.

(2) Correct reward memory and false reward memory

Under the full-attention and divided-attention conditions, the descriptive results for reward memory rates for different item types are shown in Table 2. To further analyze the effects of reward and attention condition on reward memory for different types of items, the correct reward memory rate for learned items was used as the dependent variable, and a 2 (reward condition: absent, present) × 2 (attention condition: full attention, divided attention) two-factor repeated-measures ANOVA was conducted. The analysis found a significant main effect of reward condition, $F(1, 71) = 12.45$, $p < 0.001$, $\eta_p^2 = 0.15$, 95% CI = [0.04, 0.28]; the correct memory rate for learned items with rewards ($M = 0.62$, $SD = 0.28$) was significantly higher than that for learned items without rewards ($M = 0.54$, $SD = 0.26$), indicating that participants learned rewarded items better. The main effect of attention condition was significant, $F(1, 71) = 27.87$, $p < 0.001$, $\eta_p^2 = 0.28$, 95% CI = [0.15, 0.41]; the correct memory rate for learned items in the full-attention group ($M = 0.74$, $SD = 0.35$) was significantly higher than that in the divided-attention group ($M = 0.42$, $SD = 0.36$), indicating that divided attention reduced participants' learning of the old-picture rewards. The interaction between the two was not significant, $F(1, 71) = 1.43$, $p = 0.24$.

Using the erroneous reward memory rate for critical lures as the dependent variable and reward condition and attention condition as independent variables, a two-factor repeated-measures ANOVA was conducted. The analysis found that the main effect of reward condition was not significant, $F(1, 71) = 1.88$, $p = 0.18$. The main effect of attention condition was marginally significant, $F(1, 71) = 3.56$, $p = 0.063$, $\eta_p^2 = 0.05$, 95% CI = [0.00, 0.15]; the erroneous reward memory rate for critical lures under the full-attention condition ($M = 0.40$, $SD = 0.35$) was higher than that under the divided-attention condition ($M = 0.29$, $SD = 0.36$). The interaction between reward condition and attention condition was not significant, $F < 1$.

3.2.2 Decision phase: Decision preference

Under the full-attention and divided-attention conditions, the descriptive results for decision preferences for different item types are shown in Table 3. The decision preference for learned items (i.e., selection rate) was used as the dependent variable, and a 2 (reward condition: absent, present) × 2 (attention condition: full attention, divided attention) two-factor repeated-measures ANOVA was conducted (see Figure 3A). The results showed a significant main effect of reward condition, $F(1, 71) = 53.40$, $p < 0.001$, $\eta_p^2 = 0.43$, 95% CI = [0.29, 0.55]; participants'decision preference for rewarded learned items ($M = 0.78$, $SD = 0.22$) was significantly higher than for unrewarded learned items ($M = 0.52$, $SD = 0.25$), indicating that participants were more inclined to choose learned items that had received rewards. The main effect of attention condition was significant, $F(1, 71) = 8.26$, $p = 0.005$, $\eta_p^2 = 0.10$, 95% CI = [0.02, 0.23]; the full-attention group' s decision preference for learned items ($M = 0.71$, $SD = 0.25$) was significantly higher than that of the divided-attention group ($M = 0.59$, $SD = 0.25$), indicating that divided attention reduced participants' decision preference for learned items. The interaction between the two was not significant, $F < 1$.

Table 3 Decision preferences for different item types under full-attention and divided-attention conditions ($M \pm SD$)

Reward condition	Full attention	Full attention	Divided attention	Divided attention
	Learned items	Critical lures	Learned items	Critical lures
Rewarded	0.84 ± 0.19	0.81 ± 0.19	0.72 ± 0.25	0.70 ± 0.22
Unrewarded	0.58 ± 0.27	0.47 ± 0.26	0.46 ± 0.23	0.53 ± 0.25

Using decision preference for the key lures as the dependent variable, and reward condition and attention condition as independent variables, a two-factor repeated-measures analysis of variance was conducted (see Fig. 3B). The results showed that the main effect of reward condition was significant, $F(1, 71) = 34.48$, $p < 0.001$, $\eta_p^2 = 0.33$, 95% CI = [0.19, 0.45]. Decision preference for key lures was significantly higher under the reward condition ($M = 0.75$, $SD = 0.20$) than under the no-reward condition ($M = 0.50$, $SD = 0.25$), indicating that participants were more inclined to choose key lures associated with reward-list images. The main effect of attention condition was not significant, $F < 1$. The interaction between the two was significant, $F(1, 71) = 4.11$, $p = 0.046$, $\eta_p^2 = 0.06$, 95% CI = [0.00, 0.16]. Further paired comparisons found that, under the reward condition, participants' decision preference for items in the full-attention condition was significantly higher than that in the divided-attention condition, $t(71) = 2.35$, $p = 0.021$, Cohen' s $d = 0.30$, 95% CI = [0.05, 0.56]; under the no-reward condition, there was no significant difference between the two, $t(71) = 1.08$, $p = 0.28$.

In addition, one-sample t tests comparing participants'decision preference for key lures with chance level (50%) found that, under the reward condition, decision preference for key lures in the full-attention condition was significantly higher than chance level, $t(36) = 10.15$, $p < 0.001$, Cohen' s $d = 1.67$, 95% CI = [1.33, 2.00]; in the divided-attention condition, decision preference for key lures was significantly higher than chance level, $t(35) = 5.40$, $p < 0.001$, Cohen' s $d = 0.90$, 95% CI = [0.56, 1.24]. Under the no-reward condition, participants' decision preference for key lures did not differ from chance level at either attention level: full-attention condition, $t(36) = 0.79$, $p = 0.44$; divided-attention condition, $t(35) = 0.75$, $p = 0.46$.

Figure 3. Decision preference for items under different reward conditions in the full-attention and divided-attention conditions. (A) Decision preference for learned items. (B) Decision preference for key lures.

Figure 2: Figure 3. Decision preference for items under different reward conditions in the full-attention and divided-attention conditions. (A) Decision preference for learned items. (B) Decision preference for key lures.

Fig. 3 Decision preference for items under different reward conditions in the full-attention and divided-attention conditions. (A) Decision preference for learned items. (B) Decision preference for key lures. * $p < 0.05$, n.s. indicates no significant difference.

3.2.3 Predictive analysis of reward memory on decision preference

To reveal, at the individual level, the relationship between decision preference for images and reward memory, correlation analyses were conducted between decision preference for images and, respectively, correct reward memory for rewarded learned items and incorrect reward memory for key lures under different attention conditions.

The results showed that, for learned items, whether under divided attention ($r_{(34)} = 0.47$, $p = 0.004$) or full attention ($r_{(35)} = 0.56$, p

< 0.001), correct reward memory positively predicted the decision preference for previously rewarded items; for the key lure, under both divided attention ($r_{(34)} = 0.44$, $p = 0.007$) and full attention ($r_{(35)} = 0.44$, $p = 0.006$), incorrect reward memory could positively predict the decision preference for the key lure in the rewarded condition.

A generalized linear mixed-effects model (GLMM) was further used to analyze, separately for previously learned items and key lures, whether the effect of trial-level reward memory (present vs. absent) on decision preference changed significantly under different attention conditions (full attention and divided attention). For previously learned items, participants' decision preference for rewarded learned items (yes vs. no) was taken as the dependent variable, and reward memory for that rewarded item (present vs. absent) and attention condition (full attention vs. divided attention) were entered into the model as independent variables, with each participant's random intercept and slope controlled. The results showed that the interaction between attention condition and reward memory was significant, $\beta = 3.62$, $SE = 1.03$, $z = 3.50$, $p < 0.001$, OR = 37.50, 95% CI = [4.94, 284.73]. Simple-effects analysis showed that reward memory significantly predicted decision preference under divided attention, $\beta = 2.58$, $SE = 0.83$, $z = 3.11$, $p = 0.002$; reward memory likewise significantly predicted decision preference under full attention, $\beta = 6.20$, $SE = 1.07$, $z = 5.79$, $p < 0.001$; however, divided attention significantly reduced the effect of correct reward memory on decision preference, $\beta = 2.88$, $SE = 1.01$, $z = 2.85$,

Figure 4. Decision preferences with and without reward memory under full-attention and divided-attention conditions.

Figure 3: Figure 4. Decision preferences with and without reward memory under full-attention and divided-attention conditions.

$p = 0.004$, as shown in Figure 4A.

Figure 4. Decision preferences with and without reward memory under full-attention and divided-attention conditions. (A) Decision preference for rewarded learned items predicted by correct reward memory under divided-attention and full-attention conditions. (B) Decision preference for key lures in the rewarded condition predicted by incorrect reward memory under divided-attention and full-attention conditions. Note: Error bars represent 95% CI; *** $p < 0.001$, ** $p < 0.01$.

Taking the decision preference for key lures under the rewarded and unrewarded conditions (yes vs. no) as the dependent variable, reward memory for the key lure (present vs. absent) and attention condition (full attention vs. divided attention) were entered into the model as independent variables, with each participant' s random intercept and slope controlled. The interaction effect was not significant ($\beta = 1.62$, $SE = 1.11$, $z = 1.46$, $p = 0.14$; OR = 5.06, 95% CI = [0.57, 44.70]); therefore, the interaction term was removed on the basis of this model. The results showed that the main effect of incorrect reward memory was significant, $\beta = 2.69$, $SE = 0.79$, $z = 3.43$, $p < 0.001$, OR = 14.80, 95% CI = [3.18, 68.90]; the main ⋯

The effect was also significant, $\beta = 1.01$, $SE = 0.38$, $z = 2.65$, $p = 0.008$, OR = 2.74, 95% CI = [1.30, 5.78]. The above results indicate that, regardless of whether the condition was full-attention or divided-attention, false reward memory could significantly predict decision preference, $\beta = 2.69$, $SE = 0.79$, $z = 3.43$, $p < 0.001$ (see Figure 4B). At the same time, participants in the full-attention group showed significantly higher decision preference for the key lure than those in the divided-attention group, indicating that divided attention significantly reduced participants' decision preference for the key lure.

3.3 Discussion

Experiment 2 found that divided attention significantly reduced participants' correct recognition and correct reward memory, indicating that the divided-attention manipulation could, to a certain extent, occupy participants' cognitive resources during the reward-learning phase and thereby affect their learning of the picture materials. The further reduction in decision preference also indicates that the relation between correct reward memory and subsequent decision preference is relatively sensitive to the attentional resources available during the encoding phase.

Experiment 2 focused on the influence of divided attention on false memory for the key lure and on decision preference. Although divided attention did not affect false recognition, its weakening effect on false reward memory reached marginal significance. This may be because false reward memory includes both conscious and unconscious components; divided attention occupied the conscious component of cognitive resources and had a relatively weak effect on the unconscious component. According to activation–diffusion theory, false recognition of the key lure originates from spreading activation of the learned word list, and this process is largely unconscious automatic processing. In contrast, for false reward memory, unconscious familiarity with the picture cannot enable participants to judge whether it was rewarded; therefore, conscious extraction of detailed information about whether the key lure was rewarded is still needed. This conscious component is affected by divided attention, causing false reward memory to decrease in absolute value, while a relatively high unconscious component remains. In terms of performance at the decision stage, participants in the full-attention group showed significantly higher decision preference for the key lure under the rewarded condition than those in the divided-attention group, indicating that the divided-attention manipulation could significantly reduce decision preference for the key lure. The effect of divided attention on decision preference for the key lure under the non-rewarded condition did not reach significance. Combined with the analysis of the decision-preference results in the full-attention group, this may be because decision preference for the key lure under the non-rewarded condition was already not significantly different from the 50% chance level; that is, even when the divided-attention task interfered with learning, it would only bring it closer to chance level. Because the divided-attention task mainly weakened the conscious component in memory, yet the average selection rate for the key lure under the rewarded condition in the divided-attention group was 67%, still higher than 50%, this suggests that decision preference did not completely disappear. It also suggests that decision preference may be influenced by the unconscious component of false memory and is difficult to disrupt through the divided-attention manipulation.

In summary, Experiment 2 found that false reward memory for the key lure could stably predict subsequent decision preference and might be jointly influenced by conscious and unconscious components. The conscious component could be weakened by the divided-attention manipulation, and the corresponding decision preference would also be reduced; whereas the unconscious component was retained under the divided-attention manipulation and could still predict decision preference. To further explore how the conscious and unconscious components in false memory guide value-based decision-making, and based on the similarity and covariation characteristics of false memory and true memory (Zhou Chu et al., 2007; Jou & Flores, 2013; Osth et al., 2026), with reference to previous studies (Costanzo et al.,

2021; Failes et al., 2020; Jacoby et al., 2005). Experiment 3 introduced PDP on the basis of Experiment 1 and combined it with the DRM reward-learning task paradigm. Based on the assumption in the process-dissociation model that conscious processing and unconscious processing are mutually independent, the two types of conscious components were operationally estimated to dissociate the conscious and unconscious components of false memory, and their relationships with decision-making preferences were examined.

4 Experiment 3: Effects of the Conscious and Unconscious Components of False Memory on Value-Based Decision-Making Preferences

4.1 Method

4.1.1 Participants

Using G*Power 3.0 (Faul et al., 2007), with a medium effect size set at $f = 0.2$, $\alpha = 0.05$, and statistical power $power = 0.80$, a total of 52 participants were required. In practice, 50 participants were recruited ($M = 22.58$, $SD = 1.64$, 16 males). All participants had normal or corrected-to-normal vision, participated voluntarily, and signed an informed-consent form before the formal experiment began. They received cash compensation for completing the experiment.

4.1.2 Experimental Materials

The experimental materials were the same as those in Experiment 1 and Experiment 2.

4.1.3 Experimental Design and Procedure

Experiment 3 comprised three phases (see Figure 5): the reward-learning phase, the memory-test phase, and the decision-making phase. The decision-making phase was the same as in Experiment 1 and Experiment 2. Combining the PDP and DRM paradigms, the memory test was divided into two rounds: inclusion testing and exclusion testing.

Reward-learning phase. This phase was divided into four rounds, with reference to the process-dissociation procedure (Costanzo et al., 2021; Jacoby et al., 1993). The first two rounds were Learning Phase 1, and the last two rounds were Learning Phase 2. The two learning phases were distinguished by instructional statements and background colors as cue prompts. This procedure was intended to assign each stimulus a unique, retrievable background cue, so that in the subsequent memory-test phase the "source monitoring" task could be operationalized, thereby allowing mathematical models to dissociate and estimate the conscious and unconscious memory contributions (Jacoby et al., 1993). The remaining operations were the same as in Experiment 1. The filler task was the same as in Experiment 2.

Memory Test 1: Inclusion test. The inclusion test was basically the same as the memory-test task in Experiment 1 and Experiment 2. In this phase, the test

Figure 5. Flowchart of Experiment 3

Figure 4: Figure 5. Flowchart of Experiment 3

pictures came from two learning phases: on the screen, in random order, there appeared the studied items, critical lures, and list-related items from the DRM lists corresponding to Learning Phase 1 and Learning Phase 2, as well as 15 unrelated items. Participants first needed to perform a recognition task for the pictures: press "1" for old pictures and "2" for new pictures. For old pictures, participants also needed to make a reward-source judgment: if they believed the picture was associated with reward, press "1" ; if there was no reward, press "2" ; if they could not remember, press "3." The inclusion test allowed conscious and unconscious

both memory mechanisms operate together: whether the picture material has consciously appeared during the learning phase, or whether the participant has unconsciously become very familiar with the picture material, the participant can correctly recognize old items and make affirmative responses (Jacoby et al., 1993).

Memory Test 2: Exclusion test. On the screen, the studied items, critical lures, list-related items, and 15 unrelated items from the remaining two DRM picture lists in the learning phase were presented in random order. The picture materials included in the inclusion and exclusion tests were balanced across participants. Unlike the inclusion test, in the exclusion-test phase, participants not only needed to recognize the pictures, but also had to judge the learning phase in which each old picture had appeared: if an old picture was from learning phase 1, they pressed "1" ; if an old picture was from learning phase 2, they pressed "2" ; new pictures were also assigned "2." Then, for old pictures recognized as from learning phase 1, participants performed a source-judgment task (with the same keypress requirements as in the inclusion test). By requiring participants to consciously exclude old pictures from the non-target phase, the exclusion test therefore makes it possible, under the exclusion condition, to use false alarms to "old pictures that should be excluded" and correct recognition of "old pictures from the target phase" as the basis for estimating the contributions of automatic and controlled processing, thereby separating conscious and unconscious components. In the exclusion test, participants needed to consciously identify the source of each item and deliberately exclude old items from learning phase 2. If participants failed to consciously remember the source of an item (memory failure), but the item could still elicit an unconscious response, they might mistakenly judge an old item from learning phase 2 that should have been excluded as an old item from learning phase 1 and press "1." Such erroneous judgments reflect the influence of unconsciousness on behavior when memory is not involved (Costanzo et al., 2021; Jacoby et al., 1993).

Figure 5. Flowchart of Experiment 3

Note: This flowchart is a schematic diagram. In Panel B, pictures in dashed boxes are critical lures, pictures in solid boxes are studied items, and pictures in dotted boxes are unrelated items.

Drawing on the process-dissociation procedure and calculation method of Jacoby et al. (1993), in the inclusion test, participants' recognition memory may be based on conscious components in memory, or it may be based on unconscious components. In the exclusion test, when participants cannot consciously recall the source information, they make erroneous "old" judgments (false alarms), which reflects the influence of unconsciousness. Under the exclusion condition, false judgments of "items that should be excluded" rule out the possibility of conscious recollection, thereby allowing the erroneous responses under this condition to serve as an independent basis for estimating unconscious effects.

The conscious component **R** is calculated from the difference between the probabilities of making an "old" response to studied items/critical lures in the inclusion test and in the exclusion test.

The contribution of the conscious component (denoted as R) for studied items and critical lures was calculated as follows:

$$\mathrm{R_{SI}}(\text{studied item}) = \mathrm{P}(\text{studied item} \mid \text{inclusion test}) - \mathrm{P}(\text{studied item} \mid \text{exclusion test})$$

$$\mathrm{R_{CL}}(\text{critical lure}) = \mathrm{P}(\text{critical lure} \mid \text{inclusion test}) - \mathrm{P}(\text{critical lure} \mid \text{exclusion test})$$

The contribution of the unconscious component (denoted as F) in the recognition of studied items/critical lures was calculated as follows:

$$\mathrm{F_{SI}}(\text{studied item}) = \mathrm{P}(\text{studied item} \mid \text{exclusion test})/(1 - \mathrm{R_{SI}})$$

$$\mathrm{F_{CL}}(\text{critical lure}) = \mathrm{P}(\text{critical lure} \mid \text{exclusion test})/(1 - \mathrm{R_{CL}})$$

Note: SI denotes studied item, and CL denotes critical lure.

4.2 Results

4.2.1 Memory Test Phase: Conscious and Unconscious Components

First, the memory components for studied items and critical lures under different reward conditions were analyzed. The descriptive results are shown in Table 4.

A 2 (reward condition: rewarded, unrewarded) × 2 (conscious component: unconscious, conscious) × 2 (item type: studied item, critical lure) three-factor repeated-measures analysis of variance found that the main effect of reward condition was significant, $F(1,49) = 22.25, p < 0.001, \eta_p^2 = 0.31, 95\%$ CI $= [0.14, 0.46]$; the main effect of item type was significant, $F(1,49) = 47.39, p < 0.001, \eta_p^2 = 0.49, 95\%$ CI $= [0.33, 0.62]$; the main effect of conscious component was significant, $F(1,49) = 88.49, p < 0.001, \eta_p^2 = 0.64, 95\%$ CI $= [0.51, 0.73]$; the two-factor interaction between item type and conscious component was significant, $F(1,49) = 27.71, p < 0.001, \eta_p^2 = 0.36, 95\%$ CI $= [0.19, 0.51]$; the two-factor interaction between item type and reward condition was significant, $F(1,49) = 4.03, p = 0.05, \eta_p^2 = 0.08, 95\%$ CI $= [0.00, 0.22]$; the interaction between conscious component and reward condition was not significant, $F < 1$; the three-factor interaction was significant, $F(1,49) = 9.43, p = 0.003, \eta_p^2 = 0.16, 95\%$ CI $= [0.04, 0.32]$.

Table 4. Conscious components of different item types ($M \pm SD$)

Reward condition	Conscious	Conscious	Unconscious	Unconscious
	Studied item	Critical lure	Studied item	Critical lure
Rewarded	0.33 ± 0.18	0.24 ± 0.25	0.85 ± 0.15	0.48 ± 0.46
Unrewarded	0.25 ± 0.19	0.16 ± 0.24	0.84 ± 0.17	0.33 ± 0.44

To further explore the three-factor interaction, two-factor repeated-measures analyses of variance were conducted separately for studied items and critical lures. (1) For studied items, the interaction between reward condition and conscious component was significant, $F(1,49) = 9.53, p = 0.003, \eta_p^2 = 0.16, 95\%$ CI $= [0.04, 0.32]$. Further simple-effects analysis showed that, for rewarded studied items, the difference between the conscious and unconscious components was significant, $t(49) = 25.60, p < 0.001$, Cohen' s $d = 1.37, 95\%$ CI $= [1.26, 1.48]$; for unrewarded studied items, the difference between the conscious and unconscious components was significant, $t(49) = 24.38, p < 0.001$, Cohen' s $d = 1.54, 95\%$ CI $= [1.41, 1.67]$, see Figure 6A-left. (2) For critical lures, the main effect of reward condition was significant, $F(1,49) = 15.06, p <$

0.001, $\eta_p^2 = 0.24$, 95% CI = [0.08, 0.39], with the memory effect in the rewarded condition (M = 0.36, SD = 0.39) being significantly superior to that in the unrewarded condition (M = 0.25, SD = 0.36); the main effect of awareness score was significant, $F(1,49) = 8.43$, $p = 0.006$, $\eta_p^2 = 0.15$, 95% CI = [0.03, 0.30], with the unconscious memory component (M = 0.41, SD = 0.45) being significantly higher than the conscious memory component (M = 0.20, SD = 0.24); the two-factor interaction was not significant, $F(1,49) = 3.00$, $p = 0.09$, see Figure 6A-right.

Figure 6 Results of Experiment 3. (A) The conscious and unconscious component estimates for learned items and critical lures under rewarded and unrewarded conditions; (B) decision preference under different item conditions; (C) linear regression predictions of decision preference by the unconscious and

Figure 6

Figure 5: Figure 6

conscious components for rewarded learned items; (D) linear regression predictions of decision preference by the unconscious and conscious components for critical lures under the rewarded condition. Note: Error bars represent 95% CI; *** $p < 0.001$, ** $p < 0.01$, n.s. indicates no significant difference. The scatter points in Figures 6C and 6D are the computed conscious and unconscious component results, not the raw data results of tested reward memory.

4.2.3 Decision Stage: Decision Preference

To further examine whether decision preference existed, decision preferences under different conditions were compared with the 50% chance level using one-sample t tests. The results are shown in Figure 6B. In choices between learned items, the probability that participants chose rewarded learned items ($M = 0.86$, $SD = 0.14$) was significantly higher than the 50% chance level, $t(49) = 18.59$, $p < 0.001$, Cohen' s $d = 2.63$, 95% CI = [2.35, 2.91], again verifying that participants learned the association between pictures and rewards.

In choices involving critical lures and list-related items, under the unrewarded condition participants' selection rate for critical lures ($M = 0.42$, $SD = 0.13$) was significantly lower than the 50% chance level, $t(49) = 4.66$, $p < 0.001$, Cohen' s $d = 0.66$, 95% CI = [0.37, 0.94]; under the rewarded condition, the selection rate for critical lures ($M = 0.79$, $SD = 0.16$) was significantly higher than the 50% chance level, $t(49) = 12.57$, $p < 0.001$, Cohen' s $d = 1.78$, 95% CI = [1.49, 2.06], indicating that because participants formed false memories for the critical lures during the reward-learning stage, they showed a decision preference for the critical lures under the rewarded condition and significantly avoided the critical lures under the unrewarded condition.

In the selection between learned items and list-related items, participants' selection rate for learned items without rewards ($M = 0.42, SD = 0.12$) was significantly below the 50% chance level, $t(49) = 4.92, p < 0.001$, Cohen' s $d = 0.70$, 95% CI = $[0.41, 0.98]$; the selection rate for learned items with rewards ($M = 0.82, SD = 0.13$) was significantly above the 50% chance level, $t(49) = 17.86, p < 0.001$, Cohen' s $d = 2.53$, 95% CI = $[2.24, 2.81]$, indicating that participants could distinguish, at the decision stage, whether an image was associated with a reward; participants were more inclined to select rewarded learned items and to avoid unrewarded learned items.

4.2.4 Predictive Analysis of Conscious and Unconscious Components of Memory for Decision Preferences To examine how the conscious and unconscious components in correct and false reward memories influence decision preferences for images, we conducted linear regression analyses on the conscious and unconscious components induced by different reward memories and on decision preferences for images. Linear regression models were established separately for learned items and for critical lures to test the effects of the conscious and unconscious components of recognition memory on decision preferences.

First, to predict decisions for learned items, the conscious and unconscious components of learned items were entered into the model. The results showed that the overall regression model was significant, $F(2, 47) = 7.37, p = 0.002, R^2 = 0.24$. The unconscious component significantly and positively predicted decision preference (Figure 6C-left), $\beta = 0.45, SE = 0.14, t = 3.30, p = 0.002$, 95% CI = $[0.18, 0.73]$, $R^2_{\text{partial}} = 0.19$, indicating that the unconscious component had a medium-sized effect on decision preference. By contrast, the regression coefficient for conscious recollection was not significant (Figure 6C-right), $\beta = -0.08, SE = 0.12, t = -0.64, p = 0.52$, 95% CI = $[-0.30, 0.15]$, $R^2_{\text{partial}} \approx 0.01$, suggesting that its independent influence on decision preference was relatively weak.

Second, multiple linear regression was conducted for decision preferences for critical lures, with the conscious and unconscious components of critical lures entered into the regression model. The results showed that the overall regression model was significant, $F(2, 47) = 11.27, p < 0.001, R^2 = 0.32$. The unconscious component significantly and positively predicted decision preference (see Figure 6D-left), $\beta = 0.17, SE = 0.04, t = 4.15, p < 0.001$, 95% CI = $[0.09, 0.26]$, $R^2_{\text{partial}} = 0.27$ (medium-to-large effect), indicating that the unconscious component of critical lures made a relatively large contribution to decision preference. The conscious component of false memory also positively predicted decision preference (see Figure 6D-right), $\beta = 0.16, SE = 0.08, t = 2.13, p = 0.038$, 95% CI = $[0.01, 0.32]$, $R^2_{\text{partial}} = 0.09$ (small-to-medium effect), indicating that, although the effect was relatively small, the conscious component of critical lures could also predict decision preference.

4.3 Discussion

Experiment 3 used the process-dissociation model to separate the conscious and unconscious components of false memory and separately tested their predictive effects on decision preference, thereby exploring the conscious mechanisms by which false memory influences decision preference. The results showed that, for false memories of critical lures, the unconscious component was higher than the conscious component. This explains the phenomenon observed in Experiment 2 that, after attention-diversion manipulation, the decision effect based on false memory still existed, and is consistent with the hypothesis discussed in Experiment 2 that false recognition is mainly an unconscious component and, to a certain extent, can predict decision preference.

Experiments 1 and 2 both found that reward memory can predict subsequent decision preferences, but the results of Experiment 3 indicate that this predictive effect may guide decisions through dual conscious and unconscious pathways.

Although participants could not explicitly recall the relationship between rewards and pictures during the reward-learning phase, this association may be automatically activated at the moment of decision making. More importantly, whether the memory was correct or false, the unconscious component had a stronger predictive effect on decision preference. In other words, during decision making, participants may not need to consciously recall reward details from the reward-learning phase; rather, correct choices can be made simply through the unconscious activation in the brain of the association between the picture stimulus and the reward. The unconscious component in memory is a rapid, nonintentional retrieval process that can operate when cognitive resources are lacking or when recall of contextual details fails (Jacoby et al., 1993). The present experimental results show that even when participants do not consciously recall details of the reward-learning phase, false memories may still influence value-based decisions through automatic activation of the picture–reward association.

Together with Experiment 2, it can be concluded that the influence of false memory on decision preference exists at both the conscious and unconscious levels. Specifically, unconscious activation through the spreading activation of critical lures is a prerequisite for decision making, whereas the correct conscious retrieval of the reward condition of the critical lure may be the key factor affecting the final decision; however, the influence at the unconscious level may be greater.

5 General Discussion

This study combined the DRM paradigm with a reward-learning task to systematically examine the conscious level at which false memory influences decision preferences. The results showed that: (1) participants produced false reward memories for critical lures that had not appeared, that is, they believed the critical lures had appeared and were associated with rewards; (2) participants showed decision preferences for critical lures under reward conditions, and false reward memory positively predicted decision preference; (3) the attention-distractor task significantly weakened value-based decisions guided by correct memory, whereas it had a relatively small effect on value-based decisions guided by false memory, suggesting that false-memory-guided value-based decisions contain unconscious components; (4) based on the assumption in the process-dissociation model that conscious and unconscious processing are independent of each other, operational estimates of the two types of conscious components found that the conscious and unconscious components in false memory jointly contributed to participants' decision preferences for critical lures, with the unconscious component estimated by the PDP model showing a stronger predictive effect on decision preference. From the perspectives of the behavioral effects of false reward memory, attentional-resource sensitivity, and PDP-model estimation, this study systematically answered, at the behavioral level, how false reward memory affects value-based decisions, revealing the stable role of false memory in value-based decision making and its dual-consciousness mechanism.

5.1 The Influence of False Memory on Decision Preference

This study found that false reward memory, like correct reward memory, can predict subsequent value-based decisions. Experiment 1 verified the experimental results of Wang et al. (2024), indicating that the influence of false memory on value-based decision making exists stably and is not affected by group differences. This further supports the key role of contextual memory in value-based decision making, showing that even reconstructed memories can also

guides subsequent value-based decisions. Although participants had not actually experienced obtaining rewards from critical lures during the test, when they constructed false reward memories, they were more inclined at the decision stage to select those critical-lure pictures that had never appeared before but for which they had false reward memories. Moreover, the more false reward memories participants had, the stronger their preference for critical lures, indicating that the level of false reward memory directly influenced participants' subsequent value-based decisions.

The formation of false memories stems from individuals' active construction during the reward-learning stage. According to spreading activation theory and the constructive-associative hypothesis (Collins & Loftus, 1975; Roediger et al., 2001; Wang et al., 2024), when the psychological representations of related concepts are jointly activated in a memory network, individuals may establish new conceptual associations. During the reward-learning stage, participants established an association between "studied item–reward" ; highly related concept pictures (e.g., "bread") were also semantically activated (e.g., "butter,""milk,""cream,"etc.) and thus were simultaneously linked to reward, thereby forming false reward memories. In addition, the study by Wang et al. (2024) has demonstrated that, regardless of whether there is a memory test, participants show a decision preference for critical lures during the decision stage, indicating that the "critical lure–reward" association has already been formed during the reward-learning stage. Memory-based active construction in guiding value-based decisions is highly consistent with real life: events that individuals have experienced or decisions they face are often only similar to, rather than exactly the same as, previous situations. Through memory construction, when encountering similar scenarios, the brain retrieves past experiences and extends them to new situations so as to respond more flexibly.

Another possible explanatory mechanism is fuzzy-trace theory (Brainerd & Reyna, 2002; Reyna & Brainerd, 1995), which holds that memory representations contain verbatim traces and gist traces. Among them, gist traces represent the semantic or conceptual features of input information and are fuzzy and resistant to forgetting; therefore, when facing semantically similar events that have not appeared before, extracting gist traces produces a sense of familiarity and more readily gives rise to associative false memories (Brainerd & Reyna, 2002).

When participants encoded reward-related DRM pictures during the reward-learning stage, they may have extracted the gist of the list—that is, the critical lure—and reward value became associated with the gist, thereby forming false reward memories. Therefore, in subsequent decision tasks, participants were influenced by false reward memories and preferred critical lures.

5.2 Mechanism by Which False Memories Influence Decision Preferences

The results of Experiments 2 and 3 jointly indicate that both false memories and true memories can guide value-based decisions, but the evidential bases of the two may differ. Experiment 2 found that divided attention significantly weakened value-based decisions guided by true reward memories, but this task did not affect value-based decisions guided by false reward memories. Experiment 3, based on the PDP assumption that conscious and unconscious processing are mutually independent, operationally estimated the two types of conscious components according to task performance under inclusion and exclusion conditions. The results showed that both types of conscious components positively predicted decision preferences for critical lures, and the independent explanatory contribution of the unconscious component was larger. These results suggest that value-based decisions guided by false memories do not completely depend on sufficient attentional resources, and are more likely to draw on automatically activated associative representations, with unconscious components playing a greater role in this process.

One possible explanation is that, when studied items and critical lures guide decisions, the two types of memory enter the evidential basis on which decisions rely.

are different. Existing studies have shown that divided-attention tasks impair individuals' conscious processing and source monitoring of studied items (Dewhurst et al., 2007; Otgaar et al., 2012), whereas semantic associative activation may still occur to some extent when attentional resources are limited (Zhou et al., 2007; Elliott & Brewer, 2019; Palmer et al., 2010; Turk et al., 2013). According to activation-monitoring theory and the constructive-association hypothesis (Collins & Loftus, 1975; Roediger et al., 2001; Wang et al., 2024), studied items may rely on "item–reward" bindings supported by surface-level processing, whereas critical lures are more likely to arise from deep semantic associative activation and constructive associations. It can therefore be inferred that, after a divided-attention task occupies cognitive resources, the "item–reward" binding of studied items may be more easily weakened, whereas the associative value information carried by critical lures may still influence choices to some extent; thus, under attention-limited conditions, its predictive effect is still partially retained. Therefore, although both studied items and critical lures can lead to decision preferences at the phenomenal level, they show different patterns in their sensitivity to conscious attention and in their guiding effects on value-based decisions: value-based decisions guided by correct memory rely more on reward details specific to the items, whereas those guided by false memory may rely more on automatically activated associative representations.

Furthermore, based on the PDP model, the respective contributions of conscious and unconscious components were calculated. It was found that the unconscious component had a larger independent explanatory power for subsequent decision preferences. This indicates that unconscious and conscious components may not be mutually independent: the two types of conscious components act in the same direction on decision preferences, but the relative predictive contribution of the unconscious component is greater. This result also provides a further explanation for the "value-based decisions guided by false memory under limited attention" observed in Experiment 2, suggesting that decisions guided by false memory may rely more on unconscious memory representations. According to activation-monitoring theory, false memory for critical lures is the result of cumulative, continuous unconscious activation (McDermott & Watson, 2001; Roediger, 2001). Unconscious activation and failures of its monitoring may provide critical lures with automatically transmitted reward activation (e.g., seeing the "bread" picture automatically activated the corresponding reward), thereby explaining the phenomenon that false reward memories can still influence decisions when attention is limited. Fuzzy-trace theory can also explain the present results to some extent. As a dual-process theory involving both memory and decision-making, fuzzy-trace theory proposes that verbatim traces and gist traces not only influence the formation of memory content, but also affect the retrieval of details and source information in subsequent decisions (Helm & Reyna, 2023; Reyna, 2012). Gist traces enable individuals to directly extract the "core, generalized meaning" of information and rapidly connect it with values and experiential behaviors, so that individuals can form decision preferences without fully retrieving the precise details of the items (Reyna, 2012). Moreover, a large body of research indicates that adults show a gist-trace preference and, when making decisions, rely more on gist traces representing meaning than on item-by-item weighing of precise details (Helm & Reyna, 2023; Reyna et al., 2023). Therefore, according to fuzzy-trace theory, participants' value judgments of critical lures may make greater use of low-detail, highly generalized gist-memory representations and rely less on reward details specific to the items, thereby causing decision preferences for critical lures to be driven more strongly by unconscious memory components.

Of course, the present study also found that participants' decisions were influenced by the conscious component of "critical lure–reward" memory. This is consistent with the finding of Murty et al. (2016) that memory affects decision outcomes only when participants generate retrieval at the conscious level, indicating that conscious components likewise play a role in memory-based value decisions. Therefore, in the preference for value-based decisions influenced by false memory,

in the course of improvement, dual mechanisms of conscious and unconscious processing are simultaneously present, and the unconscious mechanism has a greater predictive role.

It should be noted that the PDP estimation of conscious and unconscious components is based on the model assumption that the two types of processing are mutually independent, and this assumption may be more complex in the context of false memory. For example, according to fuzzy-trace theory, gist representations formed by semantic activation may strengthen unconscious clues and support subsequent conscious retrieval, thereby potentially producing a certain degree of coupling between the conscious and unconscious components of false memory. Even so, Jacoby and his supporters argue (Jacoby et al., 1993, 1997; Yonelinas & Jacoby, 2012) that even when statistical correlations exist, functional independence can still hold, and PDP can still be regarded as a useful framework for distinguishing conscious retrieval from unconscious automatic processing. Therefore, the present findings are only behavioral evidence obtained on the basis of the PDP independence assumption, using inclusion and exclusion tests to operationally estimate the two conscious components. In addition, the picture materials used in the present study differ from meaningless stimuli in terms of encoding cues, semantic activation, and the level of detail available for source monitoring; therefore, the relative weights of conscious and unconscious components may vary with the materials. The present conclusion is more applicable to "meaningful picture materials," whereas generalization to meaningless stimuli requires further verification. Future studies could systematically compare meaningless symbols and meaningful pictures to test whether the influence of value signals on conscious and unconscious components changes with encoding content.

6 Conclusion

This study combined the DRM paradigm with a reward-learning task and, through three experiments, revealed the conscious mechanisms by which false memory guides value-based decision-making. The results showed that false reward memory can significantly predict decision preferences for critical lures. Attentional distraction significantly weakened value-based decision-making guided by correct reward memory, but did not affect value-based decision-making guided by false reward memory, suggesting that the two types of memory-guided value-based decision-making differ in their sensitivity to attentional resources. Further, based on the assumption of the PDP model that conscious and unconscious processing are mutually independent, the two conscious components were analyzed, and it was found that both conscious and unconscious components could positively predict decision preferences for critical lures, with the unconscious component providing a larger amount of independent explanatory variance. This result provides behavioral evidence that false memory and correct memory differ in guiding value-based decision-making, and offers a new perspective for understanding how the constructiveness of memory acts on subsequent decision behavior.

References

Bao, H. W. S. (2021). bruceR: Broadly useful convenient and efficient R functions.

https://doi.org/10.32614/CRAN.package.bruceR

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bowen, H. J., Ford, J. H., Grady, C. L., & Spaniol, J. (2020). Frontostriatal functional connectivity supports reward-enhanced memory in older adults. *Neurobiology of Aging, 90*, 1-12.

Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., & Reyna, V. F. (2008). How does negative emotion cause false memories? *Psychological Science, 19*(9), 919-925.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*(5), 164-169.

Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *Quarterly Journal of Experimental Psychology, 63*(10), 2067-2079.

Chang, M., & Brainerd, C. J. (2021). Semantic and phonological false memory: A review of theory and data. *Journal of Memory and Language, 119*, 104210.

Cohen, M. S., Rissman, J., Hovhannisyan, M., Castel, A. D., & Knowlton, B. J. (2017). Free recall test experience potentiates strategy-driven effects of value on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(10), 1581-1601.

Cohen, M. S., Rissman, J., Suthana, N. A., Castel, A. D., & Knowlton, B. J. (2014). Value-based modulation of memory encoding involves strategic engagement of fronto-temporal semantic processing regions. *Cognitive, Affective, & Behavioral Neuroscience, 14*, 578-592.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Costanzo, F., Alfieri, P., Caciolo, C., Bergonzini, P., Perrino, F., Zampino, G., ···& Carlesimo, G. A. (2021). Recognition memory in Noonan syndrome. *Brain Sciences, 11*(2), 169.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17-22.

Dewhurst, S. A., Barry, C., Swannell, E. R., Holmes, S. J., & Bathurst, G. J. (2007). The effect of divided attention on false memory depends on how memory is tested. *Memory & Cognition, 35*(4), 660-667.

Doll, B. B., Shohamy, D., & Daw, N. D. (2015). Multiple memory systems as substrates for multiple decision systems. *Neurobiology of Learning and Memory, 117*, 4-13.

Elliott, B. L., & Brewer, G. A. (2019). Divided attention selectively impairs value-directed encoding. *Collabra: Psychology, 5*(1), 4.

Failes, E., Sommers, M. S., & Jacoby, L. L. (2020). Blurring past and present: Using false memory to better understand false hearing in young and older adults. *Memory & Cognition, 48*(8), 1403-1416.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Helm, R. K., & Reyna, V. F. (2023). Fuzzy trace theory: Memory and decision-making in law, medicine, and public health. In R. H. Logie, Z. Wen, S. E. Gathercole, N. Cowan, & R. W. Engle (Eds.), *Memory in science for society: There is nothing as practical as a good theory* (pp. 93-122). Oxford University Press.

Jacoby, L. L., Begg, I. M., & Toth, J. P. (1997). In defense of functional independence: Violations of assumptions underlying the process-dissociation procedure? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(2), 484-495.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General, 134*, 131-148.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: measuring recollection. *Journal of Experimental Psychology: General, 122*(2), 139-154.

Jou, J., & Flores, S. (2013). How are false memories distinguishable from true memories in the Deese-Roediger-McDermott paradigm? A review of the findings. *Psychological Research, 77*(6), 671-686.

Knowlton, B. J., & Castel, A. D. (2022). Memory and reward-based learning: A value-directed remembering perspective. *Annual Review of Psychology, 73*, 25-52.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*(3), 252-271.

Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaires compared. *Memory & Cognition, 25*, 826-837.

McDermott, K. B., & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory and Language, 45*(1), 160-176.

Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General, 145*(5), 548–558.

Osth, A. F., Zhang, L., & Williams, S. (2026). A global matching model of choice and response times in the Deese-Roediger-Mcdermott semantic and structural false recognition paradigms. *Psychological Review, 133*(4), 771–819.

Otgaar, H., Peters, M., & Howe, M. L. (2012). Dividing attention lowers children' s but increases adults' false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 204–210.

Palmer, M. A., Brewer, N., McKinnon, A. C., & Weber, N. (2010). Phenomenological reports diagnose accuracy of eyewitness identification decisions. *Acta Psychologica, 133*(2), 137–145.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*(3), 193–198.

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment and Decision Making, 7*(3), 332–359.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1–75.

Reyna, V. F., Müller, S. M., & Edelson, S. M. (2023). Critical tests of fuzzy trace theory in brain and behavior: uncertainty across time, probability, and development. *Cognitive, Affective & Behavioral Neuroscience, 23*(3), 746–772.

Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology, 69*(10), 1969–1995.

Roediger III, H. L. (2001). Spreading activation and arousal of false memories. *The nature of remembering: Essays in honor of Robert G. Crowder/American Psychological Association.*

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*(3), 385–407.

Roediger, H. L., McDermott, K. B., & Robinson, K. J. (1998). The role of associative processes in creating false memories. *Theories of memory II*, 187–245.

Schneider, E., Züst, M. A., Wuethrich, S., Schmidig, F., Klöppel, S., Wiest, R., ···& Henke, K. (2021). Larger capacity for unconscious versus conscious episodic memory. *Current Biology, 31*(16), 3551–3563.

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences, 14*, 464-472.

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition, 27*(3), 494-500.

Turk, D. J., Brady-Van den Bos, M., Collard, P., Gillespie-Smith, K., Conway, M. A., & Cunningham, S. J. (2013). Divided attention selectively impairs memory for self-relevant information. *Memory & Cognition, 41*(4), 503-510.

Umanath, S., Ries, F., & Huff, M. J. (2019). Reducing suggestibility to additive versus contradictory misinformation in younger and older adults via divided attention and/or explicit error detection. *Applied Cognitive Psychology, 33*(5), 793-805.

Wang, J., Otgaar, H., & Howe, M. L. (2024). Creating false rewarding memories guides novel decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 50*(1), 52-67.

Wang, J., Otgaar, H., Howe, M. L., & Zhou, C. (2019). A self-reference false memory effect in the DRM paradigm: evidence from eastern and western samples. *Memory & Cognition, 47*(1), 76-86.

Wang, J., Otgaar, H., Howe, M. L., Lippe, F., & Smeets, T. (2018). The nature and consequences of false memories for visual stimuli. *Journal of Memory and Language, 101*, 124-135.

Weilbächer, R. A., & Gluth, S. (2017). The interplay of hippocampus and ventromedial prefrontal cortex in memory-based decision making. *Brain Sciences, 7*(1), 4.

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science, 338*, 270-273.

Yin, X., Havelka, J., & Allen, R. J. (2024). The role of attention and verbal rehearsal in remembering more valuable item-colour binding. *Memory, 32*(9), 1158-1172.

Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition, 40*(5), 663-680.

Zhou, C., Yang, Z. L., & Qin, J. L. (2007). Are intentional processes of study list necessary for the creation of false memory: Evidence for unconscious activation. *Acta Psychologica Sinica, 39*(01), 43-49.

[Zhou Chu, Yang Zhiliang, Qin Jinliang. (2007). Does the generation of false memories depend on intentional processing of word lists: Evidence for unconscious activation. *Acta Psychologica Sinica, 39*(01), 43-49.]

The Dual-Consciousness Mechanism of False Memory Guided

Decision Making

Lu Feifei, Lu Yuetong, Zhou Chu, Wang Jianqin

(Department of Psychology, Fudan University, Shanghai 200433, China)

Abstract

The constructive nature of memory makes it susceptible to retrieval-related distortions that give rise to false memories. Whether—and how—false memories directly influence value-based decisions remains unclear. We integrated the Deese-Roediger-McDermott (DRM) paradigm with a reward-learning task to test whether false reward memories bias decision preferences via conscious or unconscious processing. Experiment 1 combined the DRM paradigm with a reward-learning task to examine the effect of false reward memories on decision preferences. Experiment 2 introduced a divided-attention task to reduce attentional resources; if false reward memories impact decision-making primarily via conscious processing, divided attention should attenuate false memory-based decision preferences. Experiment 3 employed the process-dissociation procedure (PDP) to dissociate the conscious and unconscious components of false memory. This further tested the predictive role of both components in decision preferences, clarifying mechanisms by which false memory shapes value-based choices.

We tested these hypotheses in three experiments. Experiment 1 used a within-subjects Reward Condition (Reward vs. Non-Reward; $n = 37$). Experiment 2 introduced a divided-attention task, yielding a 2 (Reward: Reward vs. Non-Reward) × 2 (Attention: Full vs. Divided) mixed design ($n = 73$). Experiment 3 used PDP to dissociate conscious and unconscious contributions of true and false memory in guiding decision-making ($n = 50$). Across experiments, during reward learning, some DRM-list items were paired with reward and others were not. In the memory phase, participants made old/new judgments and, for 'old' responses, indicated whether each picture had been rewarded. In the decision-making phase, participants maximized earnings by choosing, between two pictures, the one more likely to yield reward.

The paradigm reliably elicited false reward memories for critical lures, which consistently biased decision preferences. In Experiment 2, the divided-attention manipulation during the

encoding phase significantly reduced participants' true reward memory, and diminished preference for rewarded items. Trial-level generalized linear mixed-effects models (GLMMs) examined whether true and false reward memory could predict decision preferences. Results showed that false reward memory predicted decision preferences under both full- and divided-attention conditions. Simultaneously, decision preferences for critical lures were higher under full attention than under distraction, indicating that attentional distraction reduced participants' decision preferences. Under the PDP assumption that conscious and unconscious processes operate independently, performance in the inclusion and exclusion conditions was used to obtain operational estimates of the two components. By introducing PDP, Experiment 3 showed that false memory contained a larger unconscious component than conscious component. Models incorporating both components accounted for decision preferences for critical lures. Both conscious and unconscious components of false memory positively predicted decision preferences of critical lures, indicating complementary contributions to decision preferences.

These findings are consistent with a spreading-activation account in which reward associations propagate within semantic networks, fostering false memories that subsequently bias value-based decision-making. Together, the results show that false memory effects on decision preferences cannot be explained solely by limited attentional resources and instead reflect complementary conscious and unconscious contributions. This work clarifies how constructive memory processes shape economically relevant decisions and motivates neuroimaging research to delineate the neural pathways linking false reward memory to biased decision preference.

Keywords False memory, Decision making, Reward learning, DRM paradigm, Process-dissociation procedure

Note: Figure translations are in progress. See original paper for figures.