SCIENCE ADVANCES | RESEARCH ARTICLE

COGNITIVE NEUROSCIENCE

Salient experiences enhance mundane memories through graded prioritization

Chenyang (Leo) Lin[1]†, Wen Wen[1]†, Phillip (Xin) Cheng[1,2], Seth Schallies[1], Shrey Grover[1], Robert M. G. Reinhart[1,3,4,5]*

Salient experiences open temporal windows that boost otherwise mundane memories encoded before and after pivotal events. A proposed feature of this phenomenon is its selectivity: Salient stimuli preferentially strengthen weak memories that share semantic connections. However, evidence in humans remains inconclusive, and a key question persists: Which factors determine the presence and magnitude of memory enhancement within current neurobiological and behavioral frameworks? We present results from 10 independent studies with a total of 648 participants and provide clear evidence of both retroactive and proactive enhancements in weak memories, directly addressing ongoing debates about the existence of these effects. Notably, stronger salience learning facilitates proactive, but not retroactive, memory enhancement, challenging prevailing theories about salience's role in these processes. Instead, retroactive enhancement depends on the proximity between incidentally encoded and conditioned stimuli in a high-level feature embedding extracted from a convolutional neural network, revealing a graded prioritization mechanism. These findings offer insights into the mechanisms driving the consolidation of everyday experiences.

INTRODUCTION

Picture yourself venturing through the wilderness of Yellowstone. Suddenly, you encounter a majestic herd of bison amidst its breathtaking landscapes. Just as stumbling upon the bison imprints a strong mark on your memory, so too does the unexpected event enhance the recollection of the paths that brought you here and those leading to your next destination.

Our episodic memory prioritizes experiences that hold motivational significance or emotional salience, such as those associated with threat or reward. The prioritization occurs through encoding processes, including attention and mnemonic strategies, as well as through postencoding consolidation and retrieval mechanisms (*1*–*3*). The prioritization of memory could extend beyond salient events per se to their temporal proximities, boosting memories of neutral information preceding or occurring thereafter. This effect is known as retroactive memory enhancement (RME) and proactive memory enhancement (PME), which marks a pivotal element within adaptive memory (*4*–*6*). This capacity confers evolutionary advantages, enabling memories to adaptively incorporate motivational significance over time. The extension of memory prioritization helps develop strategies for evading predators during ancient hunts and aids modern urban navigation, for instance, by avoiding areas linked to past crimes (*7*).

A mainstream framework for understanding RME and PME is termed behavioral tagging (*8*, *9*). According to the framework (Fig. 1), the initial encoding of neutral information creates weak synaptic potentiation (a transient memory "tag" for the experience). When preceded or followed by salient events in close temporal proximity, such as fear or reward conditioning, the transient memory tag can be more effectively consolidated and transferred into long-term memory. This behavioral framework is grounded in the synaptic "tag-and-capture" neural hypothesis. On the basis of this hypothesis, the strong stimulation that (re)activates overlapping neural populations triggers the local release of plasticity-related proteins and mRNAs to be captured by the "tagged" synapses, thereby facilitating the consolidation of the weak memory traces (*8*, *10*–*12*). A central premise of the hypothesis is that, as long as the synaptic tag and the protein synthesis window overlap, the strong event can "rescue" the weak memory into long-term consolidation (*13*). However, the temporal constraints of this rescue process differ depending on event order: The synaptic tag set by a weak stimulus is relatively short-lived (approximately up to 1 hour), whereas the plasticity-related proteins generated by a strong event can remain available for 2 to 3 hours (*13*, *14*). Furthermore, the tagging and capture hypothesis suggests that the neurochemical cascade triggered by a salient event, particularly the subsequent release of plasticity-related proteins and mRNAs, is localized (*4*, *11*, *12*). Thus, weak memory tags that engage overlapping neural populations potentiated by the salient event would have high affinity for the limited plasticity-related proteins. Several human studies indicate that both RME and PME are selective for semantic categories, suggesting specificity in the behavioral tagging mechanism (*15*–*18*). Although RME and PME are often examined together within the behavioral tagging framework (*4*, *16*, *17*), a distinct aspect of PME is the altered attentional processes after salience learning. Specifically, heightened attention will continue to be allocated to stimuli that previously associated with greater motivational salience (*19*, *20*). The engagement of attentional processes during encoding in PME further distinguishes it from RME, which is primarily modulated by behavioral tagging during consolidation.

In the past decade, attempts in search of selective RME and PME have yielded mixed results. Failures to replicate the effects occurred across multiple instances, even when researchers meticulously adhered to experimental procedures of the studies that found positive results (*16*, *21*). A recent meta-analysis incorporating most existing studies cast doubt on the reliability of selective RME (*22*). Methodological differences across studies still exist, including variations in

---

[1]Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA. [2]Department of Psychology, Princeton University, Princeton, NJ, USA. [3]Center for Systems Neuroscience, Boston University, Boston, MA, USA. [4]Cognitive Neuroimaging Center, Boston University, Boston, MA, USA. [5]Center for Research in Sensory Communication and Emerging Neural Technology, Boston University, Boston, MA, USA.
*Corresponding author. Email: rmgr@bu.edu
†These authors contributed equally to this work.

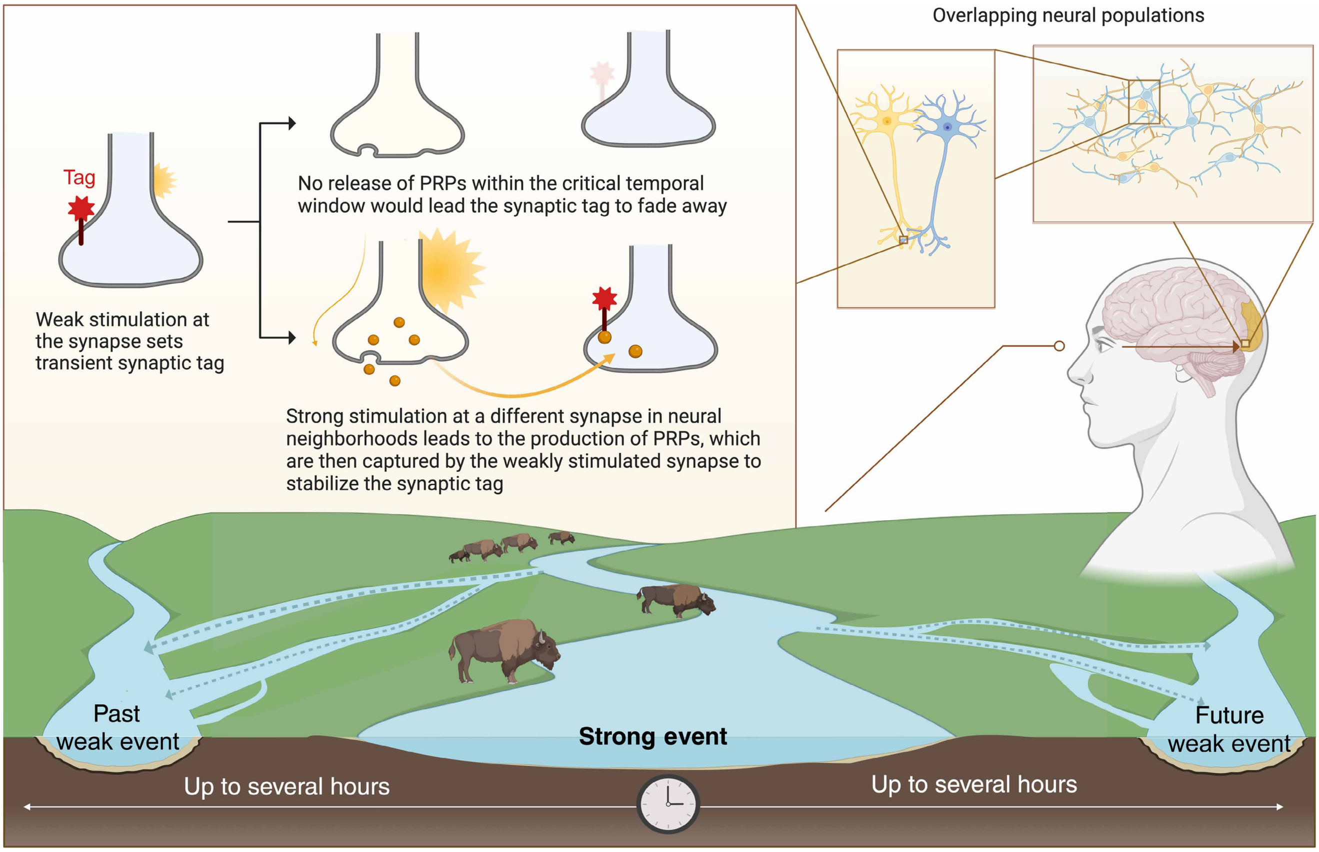

Fig. 1. Behavioral tagging hypothesis. Imagine yourself walking through the Yellowstone National Park. Then, you suddenly notice a herd of bison grazing in the nearby plains. In your excitement, you take a quick picture, eager to share the experience with your friends back home. When you reflect on this moment, the sight of those magnificent bison serves as a salient event, leaving a strong impression on your memory. Furthermore, the strong event can enhance weak memories that were encoded before and after it, as illustrated metaphorically by the water leaking from the "strong event" river into the "weak event" streams. The salient event creates a temporal penumbra, lasting several hours both before and after the event. According to the behavioral tagging hypothesis, this process involves weak memory "tags" attaching to synapses of neurons. If neighboring neurons are stimulated by a salient event, they release plasticity-related proteins that can be captured by these synaptic tags, stabilizing the memory into long-term storage. However, if there is no stimulation of neighboring neurons, these synaptic tags fade over time. The hypothesis suggests that both spatial and temporal components in neural activity are essential for strengthening initially weak memories that have occurred or are about to occur. PRPs, plasticity-related proteins.

the delay between the encoding session and the memory test, the timing between conditioning and other encoding phases, the stimuli used, and the task for incidental encoding (*16*–*18*, *23*–*26*).

Despite the contested existence of the selective memory enhancement effect, few studies have systematically examined which factors modulate the effect size. A recent work by Ouyang and Dunsmoor (*27*) identified a nonmonotonic relationship between the intensity of salience conditioning and RME, with low-intensity shocks yielding stronger retroactive enhancement, underscoring the importance of optimizing conditioning strength in behavioral tagging. Besides salience intensity, the behavioral tagging framework also emphasizes the necessity of weak memory traces: RME disappears when preconditioning items are overlearned (e.g., repeated three times) (*16*). In addition, in the seminal study by Dunsmoor *et al.* (*16*), items encoded earlier in the preconditioning phase, more temporally distant from the emotional event and likely associated with weaker memory traces before conditioning, exhibited stronger retroactive boosts. To address ongoing debates about the boundary conditions for RME and PME, large-scale datasets that encompass various experimental settings are needed (*22*).

Leveraging multiple datasets, including our own, we investigated the neurobehavioral mechanisms underlying selective memory enhancement. Our results reveal a significant salience-induced memory boost that is especially pronounced for the weak memory traces. Furthermore, PME and RME are modulated by different factors. Stronger salience learning selectively improves PME but not RME. Instead, RME is determined by the distance between incidentally encoded and conditioned stimuli in a high-level image feature embedding, showing a graded prioritization. This graded enhancement along high-level representational dimensions in memory consolidation, together with analogous effects in low-level spatial domains (*23*), points to graded memory prioritization as a general principle across representational hierarchies.

RESULTS

Specific memory enhancement in the weak memory category

In three experiments, participants underwent an encoding session followed 24 hours later by a surprise memory test (Fig. 2A). Reward

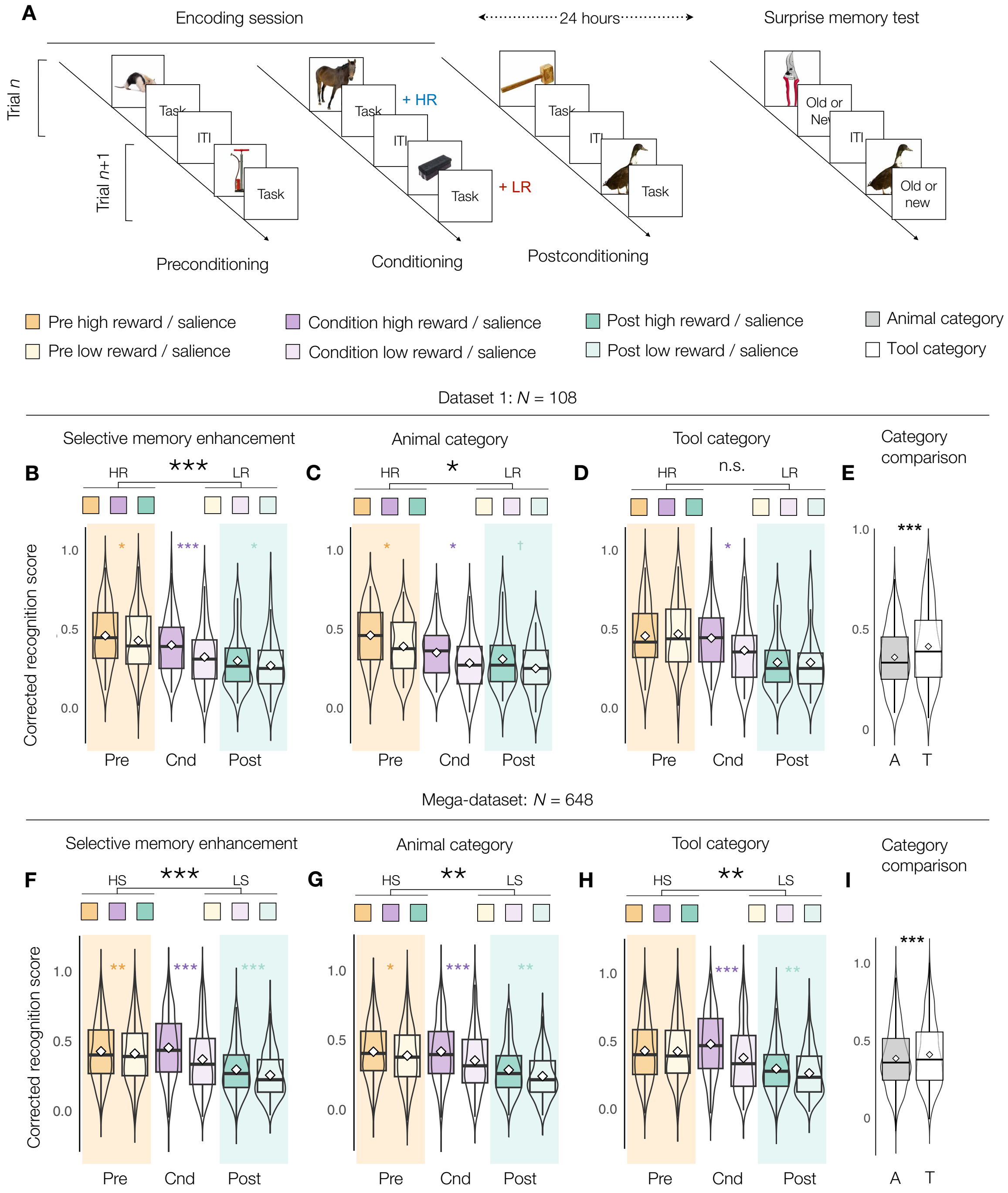

Fig. 2. Category specificity of memory enhancement. (**A**) Task in experiments 1 to 3. During the encoding session, participants either categorized images (animals or tools) or performed a match-to-sample task. In each trial, participants performed either task after the images had disappeared. During conditioning, successful task performance for one image category was associated with high reward, whereas the other category received low reward. A surprise recognition memory test occurred 24 hours later. Participants observed a mix of images from the previous day along with new images and reported whether each image was "definitely new," "maybe new," "maybe old," or "definitely old." Task, image categorization (experiments 1 and 2) or match-to-sample task (experiment 3); ITI, intertrial interval. (**B** to **E**) Results of experiments 1 to 3 (Dataset 1). HR, high reward; LR, low reward. (B) Presence of selective RME and PME in experiment 1. (C) Significant RME and PME were observed for animal images. †$P = 0.062$. (D) No significant RME or PME for tool images. (E) In Dataset 1, recognition performance for animals was significantly lower than that for tools. (**F** to **I**) Mega-dataset results. (F) Presence of selective RME and PME in the mega-dataset. (G) Salience conditioning improved the recognition performance for animals across phases. (H) The recognition performance for tools was enhanced by salience only during the conditioning and postconditioning phases. (I) Recognition performance for animals was significantly lower than that for tools in the mega-dataset. In the box-violin plots, the boxes represent interquartile ranges. Horizontal lines within boxes indicate medians, whereas diamonds represent means. The error bars show the 95% CIs. A, animal category; T, tool category; HS, high salience; LS, low salience. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; n.s., not significant.

conditioning strength and memory load were manipulated across experiments (see Materials and Methods). During the encoding session, participants either categorized images as animals or tools (experiment 1, $N = 38$; experiment 2, $N = 38$) or performed a delayed-match-to-sample task (experiment 3, $N = 32$; see Materials and Methods). The encoding session comprised three phases: preconditioning, conditioning, and, in experiments 1 and 2, a postconditioning phase. A total of 144 distinct images (48 per phase) featuring animals and tools were presented across the three phases. During the conditioning phase, one category was conditioned to high rewards and the other to low rewards. Reward assignment was counterbalanced across participants: For half, tool images were linked to high rewards, whereas for the other half, animal images were linked to high rewards. Analyses were performed on participants who correctly recalled the high-reward image category at the end of the surprise memory test on the second day. To control for individual differences in response bias, we focused on high-confidence responses and defined corrected recognition performance as the difference between the high-confidence hit rate and the high-confidence false alarm rate.

If selective RME and PME are present, images from the high-reward category encoded before and after conditioning should show higher recognition in the surprise memory test than low-reward images. Planned t tests revealed that RME was marginal in experiment 1 ($P = 0.085$) and experiment 2 ($P = 0.088$) and absent in experiment 3 ($P = 0.349$). PME emerged only in experiment 2 ($P < 0.001$) and was not detected in experiment 1 ($P = 0.659$) (Table 1 and fig. S1). These mixed outcomes, echoing prior inconsistencies in the literature, highlight the need to pinpoint the boundary conditions under which motivational salience retroactively or proactively strengthens memory.

Accordingly, we fitted a linear mixed-effects model across experiments 1 to 3 with reward magnitude (high and low), encoding phase (preconditioning, conditioning, postconditioning), and image category (animal and tool) as fixed factors (see Materials and Methods; table S1). High-reward images showed overall superior recognition [$F(1,466) = 20.619$, $P < 0.001$, $\eta_p^2 = 0.128$]. This benefit did not vary by encoding phase [reward x phase, $F(2,466) = 2.613$, $P = 0.075$, $\eta_p^2 = 0.032$]. Consistent with the main effect of reward regardless of phases, planned comparisons pooling data across experiments showed significant RME {$t(107) = 1.821$, $P = 0.036$, Cohen's $d = 0.175$, 95% confidence interval (CI) = [0.003, Inf]} and PME {$t(75) = 1.763$, $P = 0.041$, Cohen's $d = 0.202$, 95% CI = [0.002, Inf]}.

Notably, this reward-induced memory enhancement was modulated by image category [reward x phase x category, $F(2,474) = 3.337$, $P = 0.036$, $\eta_p^2 = 0.041$; Fig. 2, B to D]. Follow-up tests showed a robust reward-induced enhancement for animal images [reward main effect, $F(1,106) = 8.083$, $P = 0.005$, $\eta_p^2 = 0.105$]. This enhancement was stable across phases [reward x phase, $F(2,186) = 0.003$, $P = 0.998$, $\eta_p^2 < 0.001$]. In line with this pattern, animal images exhibited significant RME {$t(105) = 2.184$, $P = 0.016$, Cohen's $d = 0.421$, 95% CI = [0.017, Inf]} and marginal PME {$t(70) = 1.558$, $P = 0.062$, Cohen's $d = 0.359$, 95% CI = [−0.004, Inf]}.

In contrast, no significant reward-induced memory enhancement was observed for tool images [reward main effect, $F(1,105) = 0.096$, $P = 0.757$, $\eta_p^2 < 0.001$]. The reward effect varied across phases [reward x phase: $F(2,184) = 7.239$, $P < 0.001$, $\eta_p^2 = 0.225$], with memory benefits emerging only during the conditioning phase {$t(103) = 1.999$, $P = 0.024$, Cohen's $d = 0.386$, 95% CI = [0.013, Inf]}. There were no significant RME {$t(104) = -0.423$, $P = 0.663$, Cohen's $d = -0.082$, 95% CI = [−0.084, Inf]} or PME {$t(73) = 0.027$, $P = 0.489$, Cohen's $d = 0.006$, 95% CI = [−0.069, Inf]}. Thus, selective RME and PME were less evident in tool images.

In addition, a main effect of image category was found: Recognition performance for animal images was significantly lower than for tool images {$t(291) = -3.479$, $P < 0.001$, Cohen's $d = -0.204$, 95% CI = [−0.076, −0.021], two-sided paired; Fig. 2E}, suggesting that animal images are generally associated with weaker memory traces. Follow-up analyses of hit rate and false alarm revealed that this lower recognition performance was primarily driven by elevated false alarms in animal images (table S2). Compared to tool images, animal images elicited slightly lower hit rates {$t(107) = -2.061$, $P = 0.042$, Cohen's $d = -0.198$, 95% CI = [−0.053, −0.001]} and substantially higher false alarm rates {$t(107) = 4.850$, $P < 0.001$, Cohen's $d = 0.467$, 95% CI = [0.016, 0.038]}. Further analyses demonstrated that the high-level visual features of animal images were no more confusable than those of tool images (table S2). Thus, the reduced recognition accuracy for animal images likely reflects generally weaker and less reliable memory representations (*28*, *29*). For

Table 1. Ten independent experiments included in the mega-dataset. The presence or absence of selective RME and PME was analyzed on the basis of high-confidence responses only, determined with one-sided paired t tests.

Experiment	N	Type	Selective RME	Selective PME	Study type	Delay for memory test	Task type
Experiment 1	38	Reward	Absent	Absent	Lab-based	24 hours	Categorization
Experiment 2	38	Reward	Absent	Present	Lab-based	24 hours	Categorization
Experiment 3	32	Reward	Absent	Not tested	Lab-based	24 hours	Match-to-sample
Kalbe *et al.*; experiment 1 (*21*)	44	Shock	Absent	Absent	Lab-based	24 hours	Categorization
Kalbe *et al.*; experiment 2 (*21*)	80	Shock	Present	Present	Lab-based	24 hours	Categorization
Kalbe *et al.*; experiment 3 (*21*)	78	Shock	Absent	Absent	Lab-based	24 hours	Categorization
Kalbe *et al.*; experiment 4 (*21*)	83	Shock	Present	Present	Lab-based	24 hours	Categorization
Kalbe *et al.*; additional experiment (*21*)	40	Reward	Absent	Absent	Lab-based	24 hours	Match-to-sample
Sukumaran *et al.* (*30*)	170	Reward	Absent	Not tested	Online	24 hours	Match-to-sample
Hennings *et al.* (*17*)	45	Shock	Present	Present	Lab-based	24 hours	Categorization

brevity, we refer to the less reliable memories of animals as “weaker memories.” According to the behavioral tagging hypothesis (*16*), unstable memories are more likely to benefit from reward-driven consolidation mechanisms. Thus, the observed category-specific enhancement effect could be driven by the weaker memory representations of animals, making them more prioritized for stabilization through behavioral tagging.

Generalization of category-specificity effects in a mega-dataset

To evaluate the robustness and generalizability of the category-specific memory enhancement, we compiled data from 10 experiments with diverse designs, including our own, resulting in a pooled sample of 648 participants (see Table 1 for details) (*17*, *21*, *30*). Because these studies involved both aversive and rewarding conditioning paradigms, we categorized stimuli on the basis of their salience: Items associated with high rewards or shocks were classified as “high-salience,” whereas those linked to low rewards or no shocks were categorized as “low-salience.”

With the mega-dataset (table S3), we found significant RME {$t(647) = 2.468$, $P = 0.007$, Cohen’s $d = 0.097$, 95% CI = [0.005, Inf]} and PME {$t(445) = 5.442$, $P < 0.001$, Cohen’s $d = 0.258$, 95% CI = [0.028, Inf]}. Consistent with earlier findings, this dataset also revealed robust category-specific effects: A significant interaction between salience, encoding phase, and image category [$F(2,2867) = 4.895$, $P = 0.008$, $\eta_p^2 = 0.027$; Fig. 2, F to H] reinforced the idea that memory benefits are not uniform across categories. In addition, we observed an overall weaker memory of animals compared to tools [$t(1741) = -3.862$, $P < 0.001$, Cohen’s $d = -0.093$, 95% CI = [−0.034, −0.011], two-sided paired; Fig. 2I], further supporting that memory differences between categories may underlie the category-specific memory enhancement.

Further investigation of the category-specific memory enhancement in the mega-dataset showed consistent results with our data ($N = 108$). For animal images, we observed a main effect of salience [$F(1,632) = 10.131$, $P = 0.002$, $\eta_p^2 = 0.087$] and a significant salience x phase interaction [$F(2,1104) = 4.095$, $P = 0.017$, $\eta_p^2 = 0.070$]. This interaction was driven by pronounced memory differences during the conditioning phase, where high-salience animal images were remembered significantly better than low-salience ones {$t(646) = 3.673$, $P < 0.001$, Cohen’s $d = 0.289$, 95% CI = [0.035, Inf]}. Moreover, both significant RME {$t(646) = 1.758$, $P = 0.040$, Cohen’s $d = 0.138$, 95% CI = [0.002, Inf]} and PME {$t(435) = 2.874$, $P = 0.002$, Cohen’s $d = 0.273$, 95% CI = [0.019, Inf]} were observed, indicating that high salience consistently facilitated memory stabilization in the animal category.

For tool images, we also found a main effect of salience [$F(1,636) = 10.211$, $P = 0.001$, $\eta_p^2 = 0.059$] and a highly significant salience x phase interaction [$F(2,1107) = 27.780$, $P < 0.001$, $\eta_p^2 = 0.318$]. However, unlike in animals, reward-induced memory enhancement for tools was observed only during the conditioning {$t(644) = 5.389$, $P < 0.001$, Cohen’s $d = 0.424$, 95% CI = [0.071, Inf]} and the postconditioning phases [PME: $t(444) = 2.016$, $P = 0.022$, Cohen’s $d = 0.191$, 95% CI = [0.006, Inf], with no evidence of RME: {$t(646) = 0.249$, $P = 0.402$, Cohen’s $d = 0.020$, 95% CI = [−0.024, Inf]}.

In summary, our analysis of the mega-dataset derived from 10 experiments revealed consistent salience-driven memory enhancements, although the nature of these enhancements varied by category. Specifically, memories for animal images, which were generally weaker than those for tool images, showed greater benefits from retroactive memory stabilization, whereas tool memories primarily exhibited proactive enhancements.

Control analyses to assess the generalizability and robustness of category-specific memory enhancement

To further validate the category-specific memory enhancement, we conducted a series of control analyses to rule out potential confounding factors. First, we examined whether the observed category specificity was influenced by differences in conditioning effectiveness between the animal and tool categories. Conditioning effectiveness was inferred from the reward-driven enhancement of memory for images presented during the conditioning phase. We found no significant interaction between salience and stimulus category during the conditioning phase, either in our data [reward x category, $F(1,105) = 1.561$, $P = 0.214$, $\eta_p^2 = 0.010$] or in the mega-dataset [salience x category, $F(1,635) = 0.020$, $P = 0.888$, $\eta_p^2 < 0.001$], confirming that reward salience was equally effective across categories.

Next, considering the potential influence of memory load, particularly in studies lacking a postconditioning phase, we tested whether the category specificity was influenced by the presence of the postconditioning phase. Across experiments 1 to 3, we found no evidence for a four-way interaction between reward, phase, category, and memory load [$F(2,460) = 0.347$, $P = 0.556$, $\eta_p^2 = 0.002$]. In a similar analysis using the mega-dataset, we controlled for conditioning type (aversive versus reward) and memory load. The category-specific effects were not modulated by either conditioning type [salience x phase x category x conditioning type, $F(2,2800) = 0.732$, $P = 0.481$, $\eta_p^2 = 0.001$] or memory load [salience x phase x category x memory load, $F(2,2800) = 1.517$, $P = 0.218$, $\eta_p^2 = 0.001$].

Last, prior studies have shown that stimulus typicality can predict the overall item memory and that conditioning can selectively and retroactively enhance typicality ratings for related items (*17*). It is possible that the category-specific RME effect was driven by difference in typicality rather than memory strength between categories. To rule out the possibility, we compared typicality ratings between animal and tool images. No significant difference in image typicality was found between the two categories {$t(138) = 1.121$, $P = 0.264$, Cohen’s $d = 0.187$, 95% CI = [−0.157, 0.568], two-sided}, suggesting that differences in typicality do not account for the observed category-specific RME effects.

Collectively, these analyses confirm that the observed category-specific memory enhancement was primarily driven by differences in memory tag strength between the animal and tool categories, rather than by experimental variations in conditioning effectiveness, conditioning type, memory load, or image typicality.

Insights into the presence of RME and PME

Building on the robust category-specific memory enhancement, we next examined whether this pattern could offer additional insights into the presence of RME and PME in existing studies. The observed category specificity suggests two distinct scenarios for overall memory enhancement (Fig. 3). First, because only the weaker animal category exhibited RME in the mega-dataset (Fig. 2G), we anticipated that, in studies reporting RME, strong animal category effects would drive the overall result and compensate for an absence of such effects in the tool category (Scenario I). Second, because PME emerged in both the tool and animal categories, we expected that, in PME-present studies, contributions from both categories would combine to produce the overall effect (Scenario II).

Weak memories predominate in driving the overall RME

These predicted scenarios were confirmed by comparing studies with and without RME or PME (Materials and Methods). A comparison between RME-present and RME-absent studies revealed that the overall RME was primarily driven by the weaker animal memories (Scenario I). Specifically, RME-present studies demonstrated significant RME within the animal category {$t(206) = 1.925$; $P = 0.028$, Cohen's $d = 0.267$, 95% CI = [0.006, Inf]}, whereas RME-absent studies did not show this effect {$t(438) = 1.064$, $P = 0.144$, Cohen's $d = 0.101$, 95% CI = [−0.012, Inf]}. Notably, no significant RME was observed in the tool category in either RME-present or RME-absent studies {RME present: $t(199) = 1.417$, $P = 0.08$, Cohen's $d = 0.196$, 95% CI = [−0.006, Inf]; RME absent: $t(438) = -0.486$, $P = 0.686$, Cohen's $d = -0.046$, 95% CI = [−0.046, Inf]; Fig. 3, B and C}.

The presence of RME in animal but not tool category among RME-present studies should not be attributed to differences in conditioning effectiveness as conditioning was significant for both categories {animal, $t(200) = 4.196$, $P < 0.001$, Cohen's $d = 0.583$, 95% CI = [0.064, Inf]; tool, $t(187) = 6.497$, $P < 0.001$, Cohen's $d = 0.898$, 95% CI = [0.138, Inf]}. Furthermore, discrepancies between RME-present and RME-absent studies were unlikely to be driven by differences in conditioning effectiveness across studies as conditioning was successful in both study types {RME-present studies: $t(207) = 10.888$, $P < 0.001$, Cohen's $d = 0.755$, 95% CI = [0.123, Inf]; RME-absent studies: $t(439) = 5.918$, $P < 0.001$, Cohen's $d = 0.282$, 95% CI = [0.039, Inf]}.

Joint contributions of weak and strong memories underlie the overall PME

Consistent with Scenario II, PME-present studies showed significant enhancements in both animal {$t(243) = 2.181$, $P = 0.015$, Cohen's $d = 0.278$, 95% CI = [0.010, Inf]} and tool images {$t(237) = 3.179$, $P < 0.001$, Cohen's $d = 0.405$, 95% CI = [0.033, Inf]; Fig. 3E}. These

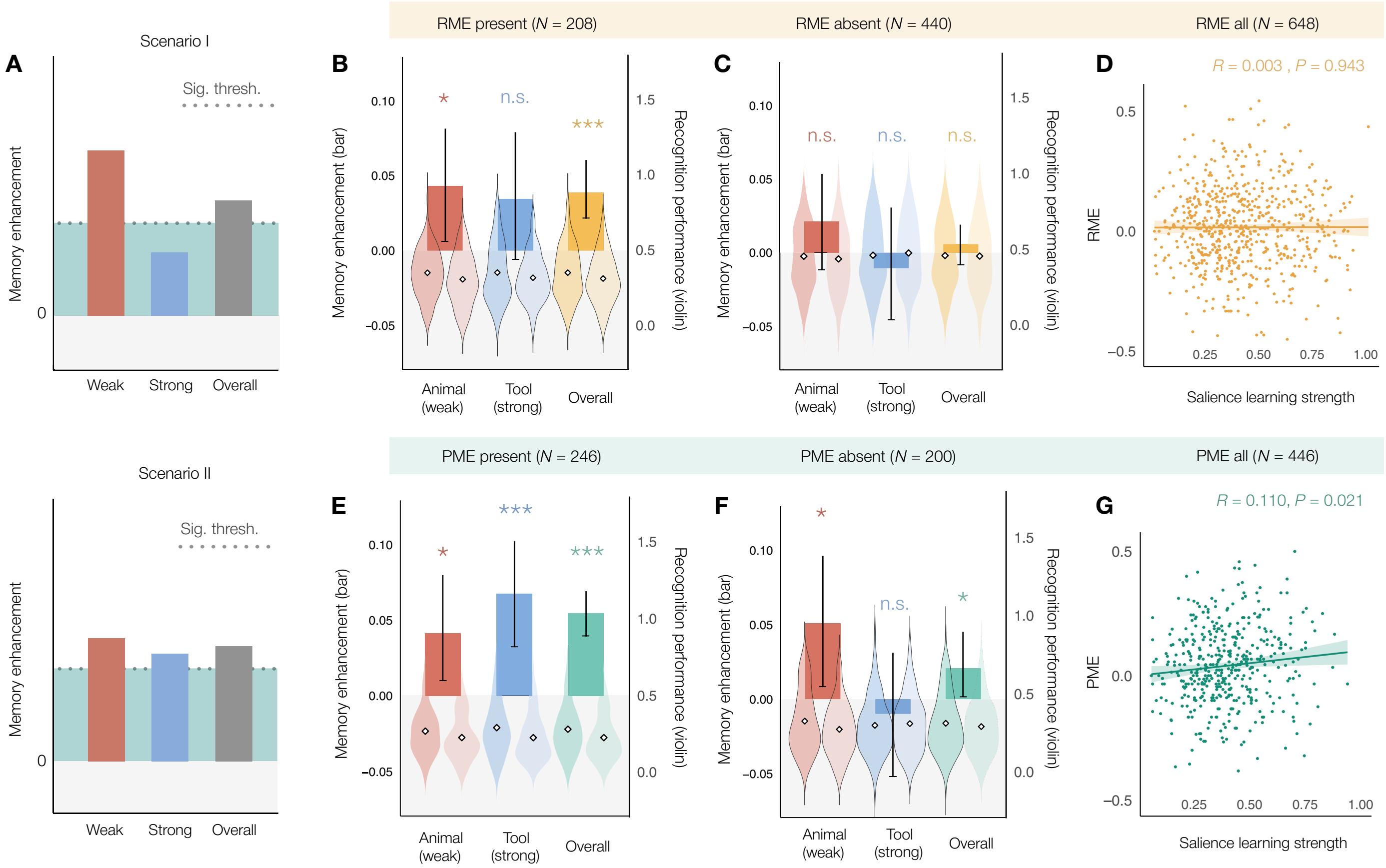

Fig. 3. Dissociable influence of memory tag strength and salience learning in RME and PME. (**A**) Two scenarios of category-specific enhancement in the mega-dataset. In Scenario I, pronounced enhancement within the weak-memory category drives the overall selective memory enhancement, despite no enhancement in the other category (top). In Scenario II, both weak and strong categories show moderate memory enhancements, jointly contributing to the overall selective memory enhancement (bottom). The dashed line marks a theoretical significance threshold. Weak: (relatively) weak memory category, animal in the present study; Strong: (relatively) strong memory category, tool in the present study. Note that animals showed weaker memories than tools in the mega-dataset, and this pattern did not differ between RME/PME present and absent studies (all $P > 0.555$). (**B** and **E**) Studies with a significant overall RME (B) conform to Scenario I: Only the weak memory category shows retroactive enhancement. Studies with significant PME (E) fit Scenario II: Both categories exhibit proactive enhancement. Violin plots (right y axis) display the distribution of recognition performance for participants receiving high versus low rewards for animals and tools; bar plots (left y axis) depict the difference between high-reward and low-reward conditions, with error bars representing 95% CI. (**C** and **F**) Combining RME absent studies (C), neither an overall RME nor RME in either image category was observed. However, when combining PME absent studies (F), the weak memory category continues to contribute to a significant overall PME, coherent with Scenario II. (**D** and **G**) The salience learning strength, measured as the average recognition performance for the conditioning phase stimuli, correlates with PME but not with RME. The shaded areas in the correlation plots represent the 95% CI.

findings suggest that salience-driven gains in both weak and strong memory traces combine to produce thee overall PME in PME-present studies. In contrast, PME-absent studies showed significant PME only in animal images {$t(189) = 1.982$, $P = 0.024$, Cohen's $d = 0.281$, 95% CI = [0.008, Inf]} not in tool images {$t(197) = -0.383$, $P = 0.649$, Cohen's $d = -0.054$, 95% CI = [−0.052, Inf]; Fig. 3F}. Because PME of the animal category persisted across study types, these results point to the tool category, whose memories are inherently stronger, as the critical determinant of whether emotional salience produces an overall PME effect.

As with RME, the differences between PME-present and PME-absent studies cannot be attributed to conditioning manipulation as conditioning was effective in both study types {PME-present studies: $t(245) = 12.497$, $P < 0.001$, Cohen's $d = 0.797$, 95% CI = [0.125, Inf]; PME-absent studies: $t(199) = 5.196$, $P < 0.001$, Cohen's $d = 0.367$, 95% CI = [0.052, Inf]}. In addition, conditioning was successful for the tool category in both PME-present {$t(220) = 7.156$, $P < 0.001$, Cohen's $d = 0.910$, 95% CI = [0.139, Inf]} and PME-absent studies {$t(198) = 3.042$, $P = 0.001$, Cohen's $d = 0.430$, 95% CI = [0.041, Inf]}, suggesting that the lack of PME in tool images in PME-absent studies is unlikely to result from ineffective conditioning.

Differential roles of attentional allocation and behavioral tagging for PME and RME

The distinct category-specific patterns in RME and PME across effect-present and effect-absent studies suggest that these effects may arise from different mechanisms. If behavioral tagging were the sole mechanism at play, RME and PME should exhibit similar patterns of category specificity, with only weaker animal memories showing significant effects. However, in PME-present studies, tool images, exhibiting stronger memories than animal images, still showed significant PME, indicating that mechanisms beyond weak memory tagging influence PME.

We hypothesized that PME is also affected by the differential allocation of attentional resources to high-salience versus low-salience categories during the postconditioning phase. The salience structure learned during conditioning may prioritize attention to high-salience categories, enhancing encoding of these stimuli compared to low-salience categories (*31*, *32*). To test this, we used mean memory performance during the conditioning phase as a proxy for the strength of salience learning. If our hypothesis is correct, better conditioning-phase memory would correlate with higher PME but not RME. We found a significant correlation between conditioning-phase memory performance and PME (Pearson's rho = 0.11, $P = 0.026$; Fig. 3G), and this effect did not differ across study types [conditioning-phase memory x PME present/absent, $F(1,442) = 0.776$, $P = 0.379$, $\eta_p^2 = 0.002$]. In contrast, no correlation was observed between salience learning strength and RME (Pearson's rho = 0.005, $P = 0.91$; Fig. 3D). Together, these findings suggest that PME is modulated by attentional prioritization of high-salience stimuli learned during conditioning, whereas RME remains independent of this attentional mechanism.

The absence of a correlation between conditioning-phase memory and RME implies that salience learning does not retroactively strengthen weaker memories uniformly but may instead operate through more complex or item-specific processes. To investigate this, we computed an item-level memory accuracy ratio (*MAR*) for each preconditioning image across experiments 1 to 3 (see Materials and Methods), defined as the proportion of high-confidence recognitions under high reward divided by that under low reward. $MAR > 1$ indicates enhancement, whereas $MAR < 1$ indicates suppression. We found that 54% of preconditioning stimuli showed $MAR > 1$ but 46% showed $MAR < 1$ (Fig. 4B). This *MAR* distribution suggests that salience-driven retroactive memory modulation can be bidirectional, enhancing some incidentally encoded memories while suppressing others.

If reward has bidirectional retroactive effects on memory, enhanced learning of the salience structure could potentially amplify both memory enhancement and suppression. To test this hypothesis, we compared the *MAR* for RME-related stimuli under two reward disparity conditions: a smaller disparity (high reward:low reward = 500:1 in experiments 1 and 3) and a larger disparity (high reward:low reward = 900:1 in experiment 2). We found a significant interaction between effect type (enhancement or suppression) and reward disparity (smaller or larger) [$F(1,92) = 5.463$, $P = 0.022$, $\eta_p^2 = 0.056$]. Specifically, a larger reward disparity was associated with trends toward stronger memory enhancement {$t(48) = 1.557$, $P = 0.063$, Cohen's $d = 0.417$, 95% CI = [−0.005, Inf]} and stronger suppression {$t(29) = -1.799$, $P = 0.041$, Cohen's $d = -0.570$, 95% CI = [−Inf, −0.004]}. No such interaction was observed for PME [$F(1,88) = 0.741$, $P = 0.392$, $\eta_p^2 = 0.008$]. These results suggest that a more prominent salience structure during conditioning amplifies both retroactive enhancement and suppression at the item level. The bidirectional influence may explain the lack of correlation between the overall RME and conditioning-phase memory performance at the subject level (Fig. 3D).

Graded memory enhancement: The role of high-level feature similarity between incidentally encoded and salient stimuli

The bidirectional effect of conditioning strength on RME at the item level suggests that RME is modulated by other factors besides salience. The behavioral tagging hypothesis suggests that memory enhancement occurs when neural populations representing weak memory traces overlap with those representing salient memories. Greater overlap facilitates the capture of plasticity-related proteins and mRNAs, enhancing the consolidation of weak memories (*4*, *6*–*8*). Thus, item-level RME may also depend on the degree of overlap between neural representations of encoded and conditioned stimuli.

We operationalized neural overlap through perceptual similarity as stimuli with similar neural representations often share high-level perceptual features, which guide sensory integration and behavior (*33*–*35*). On the basis of this, we hypothesized that weak information that shares perceptual similarities with conditioned stimuli is more likely to "capture" plasticity-related proteins released in overlapping neural populations and therefore is more likely to undergo enhancement (Hypothesis 1). Thus, we expected to observe a graded memory modulation based on the distance between incidentally encoded and conditioned stimuli in high-level visual feature space. This effect should be specific to RME (Hypothesis 2), which only relies on behavioral tagging, but not present in PME, which is primarily driven by attentional biases. We further predicted that the graded prioritization would be evident for weak memories of animal images, but not for tool images (Hypothesis 3), due to the observed category-specificity effect. As tool images exhibited overall more robust memory representations initially, there would be less opportunity for further stabilization, making them less sensitive to graded prioritization.

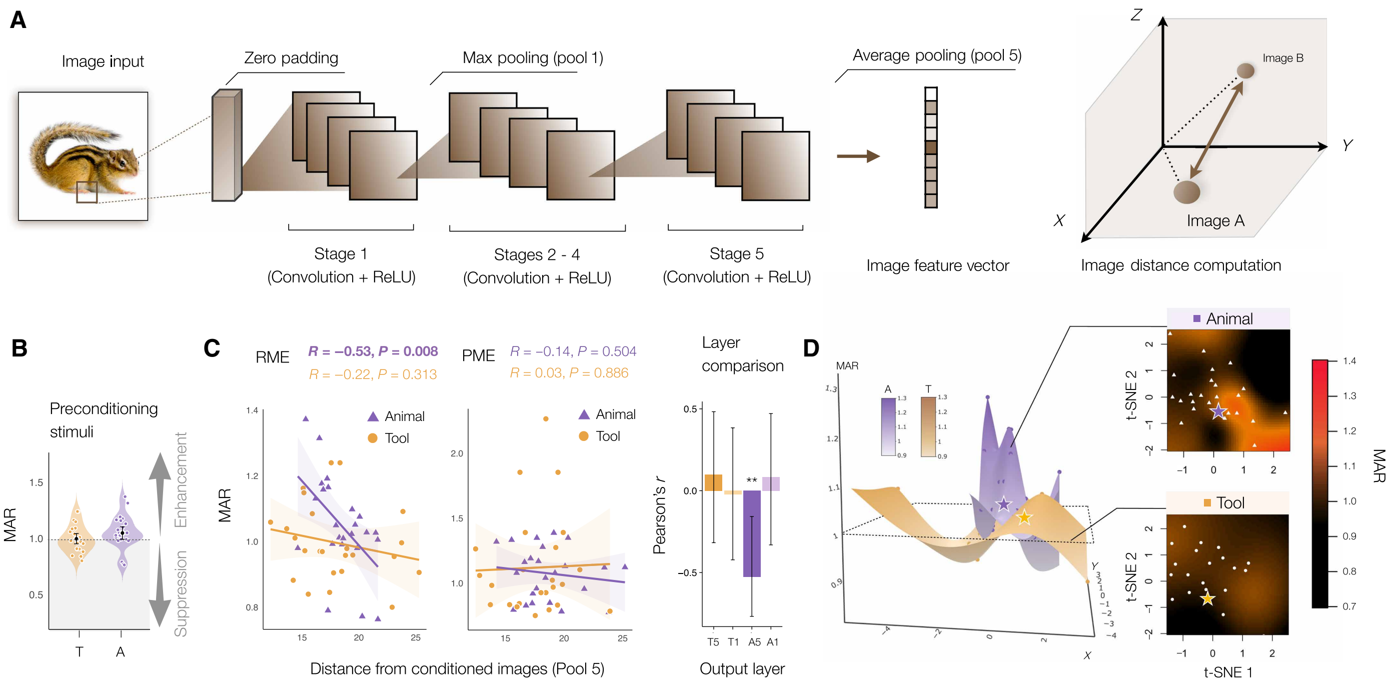

Fig. 4. Retroactive graded memory prioritization based on high-level image features. (**A**) Left: Architecture of the ResNet-18 CNN used to extract image embeddings (see Materials and Methods). Features were extracted from the global average pooling layer (Pool 5), which aggregated spatial information from preceding layers into a compact representation. Right: Illustration of computing Euclidean distances between image vectors in a simplified three-dimensional space (note that the actual feature embedding space is 512-dimensional). (**B**) Distribution of *MAR* for preconditioning stimuli across experiments 1 to 3 (A, animal images; T, tool images). (**C**) Left: *MAR* for preconditioning animal images, but not tool images, correlated with each image's distance from the centroid of high-salience images in the same category during conditioning. The *MAR* for each image was calculated by dividing the percentage of participants who correctly identified the image under the high-reward condition by the corresponding percentage of the same image under the low-reward condition. Middle: No analogous graded pattern emerges for PME. Right: Graded prioritization shows up only in Pool 5 (high-level) features, not in Pool 1 (low-level) features (see Materials and Methods). T1, tool images, output from the Pool 1 layer; T5, tool images, output from the Pool 5 layer; A1, animal images, output from the Pool 1 layer; A5, animal images, output from the Pool 5 layer. (**D**) Distribution of *MAR* for incidentally encoded images during the preconditioning phase in a t-distributed stochastic neighbor embedding (t-SNE) two-dimensional space, showing individual preconditioning images (white dots) and the centroid for all high-salience images during conditioning (stars). $^{**}P = 0.008$ (OR $P < 0.01$).

To test these hypotheses, we used convolutional neural networks (CNNs) (Fig. 4A; see Materials and Methods) to calculate the distances between incidentally encoded images and the centroid of conditioned images within each category in high-level visual feature space. These distances did not significantly differ between animal and tool categories $\{t(36) = -0.638$, $P = 0.527$, Cohen's $d = -0.184$, 95% CI = $[-2.218, 1.156]\}$. These CNN-derived distances capture each image's similarity to the conditioned stimuli in high-level visual feature space and are distinct from image typicality, as evidenced by the insignificant correlation between CNN distance and subjective typicality ratings (Pearson's rho = -0.022, $P = 0.882$). We then examined the relationship between the CNN-derived distance and *MAR* for each image in experiments 1 to 3 (Fig. 4, B and C). As predicted, we observed graded prioritization in RME (Hypothesis 1), with images farther from the centroid of conditioned images showing less enhancement [main effect of feature distance, $F(1,44) = 4.773$, $P = 0.034$, $\eta_p^2 = 0.10$; Fig. 4, C and D]. No significant correlation was observed between *MAR* and image typicality (Pearson's rho = 0.001, $P = 0.995$), suggesting that the graded prioritization is not driven by prototypicality. Supporting Hypothesis 2, no graded effect was found for PME [main effect of feature distance, $F(1,44) = 0.031$, $P = 0.861$, $\eta_p^2 < 0.101$; Fig. 4, C and D].

Further analysis revealed that retroactive graded prioritization was modulated by the strength of the memory tag [feature distance x category, $F(1,44) = 4.669$, $P = 0.036$, $\eta_p^2 = 0.10$]. Specifically, the effect occurred exclusively for the relatively weaker animal memories (Pearson's rho = -0.53, $P = 0.008$) but not for the stronger tool memories (Fig. 4C; Pearson's rho = -0.22, $P = 0.313$), thus supporting Hypothesis 3. To confirm that this graded prioritization depended on memory strength, we ran separate two-way analysis of variation (ANOVA) within each category predicting *MAR* from feature distance and memory strength (fig. S2). Within the animal category, graded prioritization was present only for weak memories (Pearson's rho = -0.896, $P < 0.001$) and was absent for strong memories [Pearson's rho = 0.402, $P = 0.195$; feature distance x memory strength: $F(1,20) = 28.72$, $P < 0.001$]. In contrast, within the tool category, neither feature distance nor memory strength showed a significant main effect, and their interaction was nonsignificant (all $P > 0.33$), potentially due to the uniformly high fidelity of tool memories. Thus, memory strength and high-level feature similarity appear to have a hierarchical influence in determining RME, and graded prioritization emerges only when the underlying memories are sufficiently fragile. Notably, the graded enhancement was only observed when using high-level visual features extracted from later layers of the CNN, rather than low-level features from earlier layers (Fig. 4C).

DISCUSSION

This study investigated how motivationally salient experiences selectively enhance otherwise mundane memories encoded close in time. Across 10 independent participant groups ($N = 648$), salience-induced enhancement was most pronounced for weak animal memories. Two key factors modulated this enhancement: the strength of salience learning and the feature similarity between incidentally encoded and salient stimuli. Specifically, salience learning strength was directly linked to PME but did not affect RME. RME, instead, was modulated by the similarity between incidentally encoded and salient stimuli, with a graded prioritization based on their distance in high-level feature space. These findings offer a comprehensive account that reconciles ongoing debates on selective memory enhancement and provides insights into the behavioral tagging hypothesis, linking long-term potentiation and protein synthesis to lasting memory formation.

One of our key findings is the category specific RME, which suggests that salience-induced memory enhancement preferentially benefits labile memory traces through behavioral tagging. Despite successful conditioning, robust memory tags are not further stabilized by salient stimulation (*8*). Consistent with this, RME has been observed in children and adolescents with attention-deficit/hyperactivity disorder (ADHD) but not in typically developing controls, who exhibit generally stronger memory performance (*36*). At a finer scale, items studied earlier in the preconditioning phase showed greater RME than those studied later, presumably because they were more temporally distant from the salient event and thus retained weaker tags at the time of strong stimulation (*16*). In our study, recognition accuracy was consistently higher for tools than for animals, likely due to tools' intrinsic utility (*37*) and their capacity to evoke motor planning unconsciously (*38*, *39*). These properties likely give rise to stronger initial memory traces for tools, rendering them less amenable to further enhancement through behavioral tagging (*8*). In contrast, animal stimuli appear to elicit more labile memory representations, making them more susceptible to postencoding modulation. Together, these results indicate that weaker memory traces are preferentially enhanced by behavioral tagging. Because tools start with a higher baseline memory strength, boosting animal memories with high salience may still leave them falling behind tools; conversely, assigning high salience to tools offers little room for further retroactive enhancement given their already strong baseline. Consequently, detecting a selective RME effect using within-subject, between-category comparisons can be challenging. Instead, the category-specific dynamic observed in our study underscores the importance of using within-category comparisons in between-subject designs to more precisely test behavioral tagging mechanisms.

Our results highlight the dissociable effects of salience structure learning, which boosts PME but not RME. In studies with overall enhancement, RME occurred exclusively for weak memories of animal stimuli (Scenario I), whereas PME was observed in both weak memories of animals and strong memories of tools (Scenario II). Memory performance during the conditioning phase correlated with PME but not with RME, suggesting that PME is driven by attentional processes beyond the behavioral tagging mechanism (*31*, *32*). The absence of a correlation between conditioning-phase memory and RME was later shown to result from the bidirectional retroactive effect of salience conditioning at the item level. The unidirectional relationship between salience structure learning and RME (*40*–*42*) has been challenged by recent evidence showing an inverted-U–shaped relationship between conditioning intensity and the magnitude of this effect (*27*). Our findings further suggest that weak memory tags can undergo both enhancement and suppression and that both effects grow stronger as the intensity of subsequent salient experiences increases. These findings challenge existing theories and emphasize the nonmonotonic effect of salience conditioning in modulating nearby memory traces.

Building on these bidirectional retroactive effects, we found that the level of enhancement was modulated by the distance between the incidentally encoded and conditioned stimuli in high-level feature space. Specifically, the degree of retroactive enhancement for weak memories correlated with their proximity to conditioned stimuli (Hypothesis 1). When memories were more distant from the conditioned stimuli, enhancement degraded. This distance-dependent gradient aligns precisely with the Glutamate Amplifies Noradrenergic Effect (GANE) model of memory consolidation (*25*, *43*, *44*), in which an emotionally salient event triggers a phasic norepinephrine surge that, when coincident with locally elevated glutamate in highly active neural assemblies, creates "hotspots" of β-adrenoceptor activation and long-term potentiation of tagged, weak traces. Outside these hotspots, lower concentrations of norepinephrine preferentially engage α-1/α-2 receptors to suppress less relevant representations via lateral inhibition (*43*, *45*). Our data mirrored this architecture: weak memories positioned within the neural "hotspot" received robust stabilization, whereas those situated farther away failed to engage β-adrenoceptor–mediated potentiation and were correspondingly less enhanced. In addition, the graded prioritization was only observed for the weak memories of animal stimuli but not tools (Hypothesis 3), likely because memories for tools are more robust and less susceptible to further enhancement through salience-induced plasticity. Last, this effect was evident only for stimuli encoded before conditioning, not for those encoded afterward (Hypothesis 2), as RME is primarily mediated by behavioral tagging mechanisms, whereas PME is further influenced by attentional bias. Collectively, these findings provide insights into the neurobehavioral mechanisms underlying salience-induced memory enhancement, with high-level visual similarity between incidentally encoded and salient stimuli being a key determinant (*4*, *12*, *46*).

Our finding that high-level feature proximity critically influences memory prioritization extends prior works on physical and temporal proximity. Whereas earlier studies demonstrated that reward can retroactively enhance memory via mechanisms linked to spatial and temporal distances—presumably through dopaminergic modulation and hippocampal place cell replays (*23*)—our work shows that salient events also selectively prioritize memories on the basis of high-level visual similarity. This suggests that dopaminergic modulation of hippocampal circuits not only supports spatial navigation but also facilitates the construction of abstract, nonspatial cognitive maps. Such graded prioritization likely underpins the formation of high-level mental models that guide adaptive behaviors like reward-seeking and punishment avoidance (*47*–*49*).

This study opens several avenues for future research. First, using psychophysiological and neurophysiological measures to directly assess conditioning strength could clarify its role in both RME and PME. Second, investigating how similarity in different representational spaces (e.g., sensory modalities or semantic features) influences memory enhancement would shed light on how memory tags integrate across levels. For instance, does a single memory tag for an object encompass information across both low-level and high-level features, or do multiple tags exist, each representing a distinct aspect of the object? Third, future studies could examine the neural mechanisms of behavioral tagging using neuroimaging techniques to measure neural representational similarities between weak memories and salient stimuli, potentially identifying neural markers for graded memory enhancement. Last, building on our multiexperiment approach, future research should aim to recruit larger, more diverse samples, where task parameters and stimulus sets are systematically manipulated to test the generality and boundary conditions of the observed memory enhancement effects.

As a final reflection, consider again the powerful moment of encountering a bison in Yellowstone. Although your conversations with fellow travelers are memorable, it is the seemingly ordinary sights and sounds, like a cottontail rabbit along the way, that are unexpectedly enhanced in memory. These neutral stimuli, which share visual similarities with the bison, benefit from enhanced consolidation, illustrating how weak memories can be selectively prioritized on the basis of their similarity to salient events.

MATERIALS AND METHODS

Participants

A total of 123 participants signed up for experiments 1 to 3. All participants reported having no metal implants in the head, no history of neurological problems or head injury, no current use of psychoactive medication, normal color vision, and normal or corrected-to-normal vision and hearing. Power analyses were conducted on the basis of preliminary data from each experiment to estimate the required sample size. Specifically, effect sizes were calculated using partial eta-squared (η_p^2) on the basis of the preconditioning reward main effect. The effect size estimates ranged from $\eta_p^2 = 0.40$ to 0.43. The estimates were then used to determine the minimum number of participants required to achieve a statistical power of 0.8 while controlling for a type I error rate at an alpha level of 0.05. The analyses indicated a minimum sample size of 25 to 28 participants.

Participants performed a categorization task in experiments 1 and 2 and a delayed match-to-sample task in experiment 3. The high-to-low reward ratio was 500:1 in experiments 1 and 3 and 900:1 in experiment 2. Six participants in experiment 2 and one participant in experiment 3 who reported that the recognition memory test was not a surprise or failed to recall the correct image category associated with the high reward were excluded from further analyses. To ensure balanced memory performance between participants rewarded for the tool and animal categories, four participants who did not respond to the reward manipulation were excluded (using a cutoff at the fifth percentile, based on $z = 2$). These exclusions applied to both experiments 1a and 1c. This process ensured comparability of memory performance across groups. The final sample consisted of 108 participants: experiment 1 ($N = 38$, 16 females, mean age $= 19.64 \pm 5.07$ years); experiment 2 ($N = 32$, 24 females, mean age $= 19.81 \pm 2.52$ years); and experiment 3 ($N = 38$, 24 females, mean age $= 18.82 \pm 1.20$ years). The study was approved by the Boston University Ethics Committee (no. 4230E), and all participants provided written informed consent before viewing any study materials.

Paradigm and stimuli

Experiment 1 consisted of two sessions: an incidental encoding session followed by a surprise recognition memory test. The incidental encoding session was divided into three phases: preconditioning, reward conditioning, and postconditioning. The images used in this study were graciously provided by the authors of (*16*). The image typicality was controlled across encoding phases and categories (*16*, *21*). A two-way ANOVA with Phase (preconditioning, conditioning, and postconditioning) and Category (animal and tool) as between-subjects factors on mean typicality ratings confirmed that there was no significant main effect [Phase: $F(2,138) = 0.17$, $P = 0.844$; Category: $F(1,138) = 1.22$, $P = 0.271$] or Phase × Category interaction [$F(2,138) = 0.001$, $P = 0.999$]. All stimulus sizes were standardized to 500 × 500 pixels. The order of image stimuli within each encoding phase was randomized across participants. During the conditioning phases, each image stimulus was presented for 2 s, followed by a stimulus-response interval of 4 ± 2 s, during which a fixation cross was displayed on a blank background. This was then followed by an intertrial interval of 0.5 s. During the encoding phases, participants made two-alternative forced-choice picture categorizations ("animal" or "tool"). After each stimulus presentation, participants were asked to classify the picture as either an animal or a tool by pressing the "F" or "J" button on a keyboard. The corresponding response buttons were counterbalanced across participants, and responses timed out after 5 s. The conditioning phase occurred immediately after the preconditioning phase, with participants completing the same task as in the preconditioning phase. However, after each trial, participants received feedback on whether they had categorized the image correctly. If correct, they were rewarded with either 1 or 500 points, depending on the stimulus category (animal or tool), with the high-reward category counterbalanced across participants. Participants were informed that the points would translate into a cash bonus at the end of the session. The postconditioning phase followed immediately after the conditioning phase and was identical to the preconditioning phase. Different images were presented in each encoding phase. Each phase included 24 animal images and 24 tool images, with an ~5-min interval between phases. Experiment 2 had an identical structure to experiment 1, with the key difference being a larger reward disparity during conditioning (1 versus 900 points). This increased disparity was designed to amplify the strength of reward conditioning in experiment 2.

Experiment 3 followed a similar structure to experiment 1, with a few differences. First, the reward conditioning phase in experiment 3 replaced the point rewards (1 or 500 points) with monetary rewards (1¢ or $5). Participants were informed that their total cash reward would be divided by eight and provided to them at the end of the second-day session to emphasize a clearer reward representation. Second, instead of the image categorization task, participants performed a delayed match-to-sample task. The task involved presenting a target image for 2 s, followed by a 4 ± 2 s delay, and then displaying both the target and a foil image side-by-side. Participants had 5 s to respond by pressing "F" or 'J" on the keyboard to select the matching image on the left or right, respectively. As in experiment 1, feedback indicated whether their response was correct or incorrect. During the reward conditioning phase, an additional message informed participants of their monetary reward after each trial. In addition, experiment 3 only included preconditioning and conditioning phases to specifically target the retroactive enhancement effect. Participants showed high accuracies in the categorization task ($M = 0.99$, $SE = 0.006$) and delayed match-to-sample task ($M = 0.99$, $SE = 0.002$), indicating strong task engagement. The primary aim of experiment 3 was to remove the postconditioning phase to reduce the overall memory load and assess how this manipulation modulates reward-related memory enhancement.

The surprise recognition memory test for all experiments took place 24 hours after the incidental encoding session. Upon arrival, participants were asked whether they had anticipated a memory test for the images they had seen the previous day. In experiments 1 and 2, the test comprised 72 high-reward targets (CS+), 72 low-reward targets (CS−), 72 novel animal lures, and 72 novel tool lures, whereas in experiment 3, it consisted of 48 CS+, 48 CS− targets, 48 novel animal lures, and 48 novel tool lures. During the memory test, participants classified images as either old (from the previous day) or new (not seen before). They reported their confidence using four options: definitely old, maybe old, maybe new, or definitely new. Each image was presented for 2 s, followed immediately by the four response options. Responses were self-paced, with participants encouraged to prioritize accuracy. To maintain participant engagement, feedback was provided after each response indicating whether the participant correctly identified the image as old or new.

Mega-dataset

A total of 648 participants from 10 separate experiments (including three of our own above) were pooled for the mega-analysis. Participants in seven experiments ($N = 412$) underwent selective PME testing. Five experiments used electric shocks for conditioning ($N = 330$), and the remaining four experiments used reward for conditioning ($N = 284$). Detailed information can be found in Table 1.

The 10 experiments in the mega-dataset were reclassified as RME-present or PME-present or RME-absent or PME-absent on the basis of one-tailed *t* tests. This reclassification established a common standard across experiments to assess whether the observed category-specific effect reconciles prior inconsistencies in detecting RME and PME across studies.

Statistical analyses

Category-specific memory enhancement

Our analyses focused on high-confidence responses only (definitely old and definitely new). Recognition memory was measured using corrected recognition, calculated as the hit rate minus the false alarm rate to control for individual differences in response bias. A previous imaging study has supported the neuroanatomical separation between memory recollection and familiarity (*50*). Emotional salience specifically amplifies recollection, as evidenced by heightened amygdaloid responses during encoding for later recollected stimuli, compared to familiar or forgotten ones (*51*). In addition, a recent work with primate models found no impairment in familiarity-based item recognition following amygdaloid lesions (*52*). Therefore, high-confidence responses in recognition memory ("definitely old" and "definitely new"), indicating recollection rather than mere familiarity, were used for testing emotional memory consolidation across all experiments.

To assess salience-induced memory enhancement effects within specific stimulus categories, we conducted independent *t* tests to compare memory performance between high-reward and low-reward associations within each category (between-subjects). Note that one-tailed *t* tests were used solely as post hoc analyses following significant main or interaction effects, not as standalone hypothesis tests. Again, memory performance for comparison was assessed using corrected recognition, calculated within the corresponding category. One-sided *t* tests were used because prior research consistently shows that RME and PME enhance memory, not impair it (*22*). Given this established pattern, any unexpected results suggesting a negative effect of high rewards are not only improbable but irrelevant to our hypothesis. Moreover, because of the typically small effect sizes in this area, one-sided *t* tests increase statistical power by focusing on the predicted direction and reduce the risk of type II errors.

To investigate the category-specific memory enhancement, we applied a linear mixed-effects model using reward magnitude (high and low), encoding phase (preconditioning, conditioning, and postconditioning), and image category (animal and tool) as fixed effects. Participant group and individual participants were treated as random effects to account for repeated measures across participants and experiments, structured as follows

$$\begin{aligned}&\text{High confidence corrected recognition score}\\&\sim \text{Reward magnitude} \times \text{Encoding phase} \times \text{Image category}\\&+(1 \mid \text{Study/Participant})\end{aligned}$$

Analyses were conducted using the "lme4" package in RStudio Version 2023.06.2 + 561. Degrees of freedom for fixed effects were estimated using the Satterthwaite approximation. The analysis revealed a significant three-way interaction among reward magnitude, encoding phase, and image category. We subsequently examined recognition performance for animal and tool images when they were categorized under high-reward versus low-reward conditions using independent *t* tests.

Between-experiment comparison

To investigate the dissociable roles of memory tag strength in selective RME and PME, we categorized experiments on the basis of whether they demonstrated selective RME and PME. Positive selective RME and PME effects were identified using paired *t* tests. Considering the observed category-specific memory enhancement, we developed strong hypotheses regarding each category's contribution to the overall selective memory enhancement. Specifically, we posited that either the weak memory category (animals) contributes alone to the overall selective memory enhancement or both the weak (animals) and strong (tools) categories contribute.

We conducted paired *t* tests on the overall selective memory enhancement for each study group type and independent *t* tests for animal and tool images under high-reward versus low-reward conditions. In addition, we performed Pearson correlation tests between the conditioning strength, measured by the average recognition performance for images presented during the conditioning phase, and RME and PME.

Calculating MARs

The *MAR* for each stimulus was calculated as the ratio of participants accurately recognizing the stimulus with high confidence in the high-reward condition to those in the low-reward condition during the surprise recognition test. For example, if 16 of 20 participants (80%) recognized a given image in the high-reward condition and 12 of 20 (60%) did so in the low-reward condition, the *MAR* for that image is 80%/60% = 1.33.

Extracting high-level features using a CNN

To extract high-level image features, we used the Deep Learning Toolbox in MATLAB (MathWorks, Natick, MA), using ResNet-18, a CNN with 18 layers. Trained on over a million images, this network can categorize images into 1000 object classes. Thus, ResNet-18 has learned sophisticated feature representations across a diverse array of images. We resized all input images to 224 by 224 by 3 before feeding them into the network. The network constructs a hierarchical representation of the input images, with deeper layers generating higher-level features based on the lower-level features extracted in earlier layers. To obtain feature representations, we extracted outputs from the "pool5" layer, which is the global pooling layer aggregating higher-level features from all spatial locations of the images. This provided us with 512 features for each image. Each image was then mapped to a 512-dimensional space, so distances could be calculated among the images in this high-dimensional space as a measure of similarity in high-level image features.

For animal and tool images presented during the conditioning phase, we calculated an average image feature vector on the basis of all images shown during that phase for each category separately. We then computed the distance between each incidentally encoded image (preconditioning and postconditioning) and the average image vector of the same category from the conditioning phase. This distance served as a measure of similarity to the conditioned stimuli. We then calculated the *MAR* for each incidentally encoded image by dividing the percentage of individuals who successfully recognized the image with high confidence under the high-reward condition by the percentage under the low-reward condition. Pearson tests were then performed to examine the correlation between the distance to the conditioned stimulus mean and the *MAR*.

Supplementary Materials

This PDF file includes:
Figs. S1 and S2
Tables S1 to S4
References

REFERENCES AND NOTES

1. S. Hamann, Cognitive and neural mechanisms of emotional memory. *Trends Cogn. Sci.* **5**, 394–400 (2001).
2. K. S. LaBar, R. Cabeza, Cognitive neuroscience of emotional memory. *Nat. Rev. Neurosci.* **7**, 54–64 (2006).
3. D. Shohamy, R. A. Adcock, Dopamine and adaptive memory. *Trends Cogn. Sci.* **14**, 464–472 (2010).
4. J. E. Dunsmoor, V. P. Murty, D. Clewett, E. A. Phelps, L. Davachi, Tag and capture: How salient experiences target and rescue nearby events in memory. *Trends Cogn. Sci.* **26**, 782–795 (2022).
5. K. Murayama, C. Kuhbandner, Money enhances memory consolidation—But only for boring material. *Cognition* **119**, 120–124 (2011).
6. Y. Zhu, Y. Zeng, J. Ren, L. Zhang, C. Chen, G. Fernandez, S. Qin, Emotional learning retroactively promotes memory integration through rapid neural reactivation and reorganization. *Elife* **11**, e60190 (2022).
7. D. Moncada, F. Ballarini, H. Viola, Behavioral tagging: A translation of the synaptic tagging and capture hypothesis. *Neural Plast.* **2015**, 650780 (2015).
8. F. Ballarini, D. Moncada, M. C. Martinez, N. Alen, H. Viola, Behavioral tagging is a general mechanism of long-term memory formation. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14599–14604 (2009).
9. T. Rogerson, D. J. Cai, A. Frank, Y. Sano, J. Shobe, M. F. Lopez-Aranda, A. J. Silva, Synaptic tagging during memory allocation. *Nat. Rev. Neurosci.* **15**, 157–169 (2014).
10. C. H. Bailey, M. Giustetto, Y. Y. Huang, R. D. Hawkins, E. R. Kandel, Is heterosynaptic modulation essential for stabilizing Hebbian plasticity and memory? *Nat. Rev. Neurosci.* **1**, 11–20 (2000).
11. U. Frey, R. G. Morris, Synaptic tagging and long-term potentiation. *Nature* **385**, 533–536 (1997).
12. R. L. Redondo, R. G. M. Morris, Making memories last: The synaptic tagging and capture hypothesis. *Nat. Rev. Neurosci.* **12**, 17–30 (2011).
13. U. Frey, R. G. Morris, Weak before strong: Dissociating synaptic tagging and plasticity-factor accounts of late-LTP. *Neuropharmacology* **37**, 545–552 (1998).
14. A. Govindarajan, I. Israely, S.-Y. Huang, S. Tonegawa, The dendritic branch is the preferred integrative unit for protein synthesis-dependent LTP. *Neuron* **69**, 132–146 (2011).
15. D. Clewett, J. Dunsmoor, Novel strategies for expanding memory's penumbra in aging. *Trends Cogn. Sci.* **27**, 120–121 (2023).
16. J. E. Dunsmoor, V. P. Murty, L. Davachi, E. A. Phelps, Emotional learning selectively and retroactively strengthens memories for related events. *Nature* **520**, 345–348 (2015).
17. A. C. Hennings, J. A. Lewis-Peacock, J. E. Dunsmoor, Emotional learning retroactively enhances item memory but distorts source attribution. *Learn. Mem.* **28**, 178–186 (2021).
18. A. Patil, V. P. Murty, J. E. Dunsmoor, E. A. Phelps, L. Davachi, Reward retroactively enhances memory consolidation for related items. *Learn. Mem.* **24**, 65–69 (2017).
19. A. J. Kim, B. A. Anderson, Threat reduces value-driven but not salience-driven attentional capture. *Emotion* **20**, 874–889 (2020).
20. B. A. Anderson, P. A. Laurent, S. Yantis, Value-driven attentional capture. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10367–10371 (2011).
21. F. Kalbe, L. Schwabe, On the search for a selective and retroactive strengthening of memory: Is there evidence for category-specific behavioral tagging? *J. Exp. Psychol. Gen.* **151**, 263–284 (2022).
22. D. Koevoet, A. Postma, Is there selective retroactive memory enhancement in humans?: A meta-analysis. *Psychon. Bull. Rev.* **31**, 531–540 (2024).
23. E. K. Braun, G. E. Wimmer, D. Shohamy, Retroactive and graded prioritization of memory by reward. *Nat. Commun.* **9**, 4886 (2018).
24. L. Bréchet, S. B. Hausmann, R. Mange, B. Herbelin, O. Blanke, A. Serino, Subjective feeling of re-experiencing past events using immersive virtual reality prevents a loss of episodic memory. *Brain Behav.* **10**, e01571 (2020).
25. D. Clewett, J. Dunsmoor, S. L. Bachman, E. A. Phelps, L. Davachi, Survival of the salient: Aversive learning rescues otherwise forgettable memories via neural reactivation and post-encoding hippocampal connectivity. *Neurobiol. Learn. Mem.* **187**, 107572 (2022).
26. J. P. Oyarzún, P. A. Packard, R. de Diego-Balaguer, L. Fuentemilla, Motivated encoding selectively promotes memory for future inconsequential semantically-related events. *Neurobiol. Learn. Mem.* **133**, 1–6 (2016).
27. L. Ouyang, J. E. Dunsmoor, Emotional intensity produces a linear relationship on conditioned learning but an inverted U-shaped effect on episodic memory. *Learn. Mem.* **31**, a053982 (2024).
28. M. Singer, Strength-based criterion shifts in recognition memory. *Mem. Cognit.* **37**, 976–984 (2009).
29. M. Singer, J. T. Wixted, Effect of delay on recognition decisions: Evidence for a criterion shift. *Mem. Cognit.* **34**, 125–137 (2006).
30. P. Sukumaran, N. Kazanina, C. Houghton, Reward conditioning may not have an effect on category-specific memory. *Sci. Rep.* **13**, 22297 (2023).
31. R. A. Adcock, A. Thangavel, S. Whitfield-Gabrieli, B. Knutson, J. D. E. Gabrieli, Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron* **50**, 507–517 (2006).
32. B. C. Wittmann, N. Bunzeck, R. J. Dolan, E. Düzel, Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *Neuroimage* **38**, 194–202 (2007).
33. Z. Kourtzi, N. Kanwisher, Representation of perceived object shape by the human lateral occipital complex. *Science* **293**, 1506–1509 (2001).
34. L. K. Tyler, H. E. Moss, M. R. Durrant-Peatfield, J. P. Levy, Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain Lang.* **75**, 195–231 (2000).
35. E. Yee, D. M. Drucker, S. L. Thompson-Schill, fMRI-adaptation evidence of overlapping neural representations for objects related in function or manipulation. *Neuroimage* **50**, 753–763 (2010).
36. V. Baumann, T. Birnbaum, C. Breitling-Ziegler, J. Tegelbeckers, J. Dambacher, E. Edelmann, J. R. Bergado-Acosta, H.-H. Flechtner, K. Krauel, Exploration of a novel virtual environment improves memory consolidation in ADHD. *Sci. Rep.* **10**, 21453 (2020).
37. M. Platt, C. Padoa-Schioppa, "Chapter 29—Neuronal representations of value" in *Neuroeconomics*, P. W. Glimcher, C. F. Camerer, E. Fehr, R. A. Poldrack, Eds. (Academic Press, 2009), pp. 441–462.
38. M. S. Beauchamp, K. E. Lee, J. V. Haxby, A. Martin, Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* **34**, 149–159 (2002).
39. S. H. Creem-Regehr, J. N. Lee, Neural representations of graspable objects: Are tools special? *Brain Res. Cogn. Brain Res.* **22**, 457–469 (2005).
40. J. de Carvalho Myskiw, F. Benetti, I. Izquierdo, Behavioral tagging of extinction learning. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1071–1076 (2013).
41. P. Lopes da Cunha, R. Tintorelli, J. Correa, P. Budriesi, H. Viola, Behavioral tagging as a mechanism for aversive-memory formation under acute stress. *Eur. J. Neurosci.* **55**, 2651–2665 (2022).
42. S. Vishnoi, S. Raisuddin, S. Parvez, Behavioral tagging: A novel model for studying long-term memory. *Neurosci. Biobehav. Rev.* **68**, 361–369 (2016).
43. M. Mather, D. Clewett, M. Sakaki, C. W. Harley, Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behav. Brain Sci.* **39**, e200 (2016).
44. G. R. Poe, S. Foote, O. Eschenko, J. P. Johansen, S. Bouret, G. Aston-Jones, C. W. Harley, D. Manahan-Vaughan, D. Weinshenker, R. Valentino, C. Berridge, D. J. Chandler, B. Waterhouse, S. J. Sara, Locus coeruleus: A new look at the blue spot. *Nat. Rev. Neurosci.* **21**, 644–659 (2020).

45. C. G. Abdallah, L. A. Averill, J. H. Krystal, S. M. Southwick, A. F. T. Arnsten, Glutamate and norepinephrine interaction: Relevance to higher cognitive operations and psychopathology. *Behav. Brain Sci.* **39**, e201 (2016).
46. M. Nomoto, N. Ohkawa, H. Nishizono, J. Yokose, A. Suzuki, M. Matsuo, S. Tsujimura, Y. Takahashi, M. Nagase, A. M. Watabe, F. Kato, K. Inokuchi, Cellular tagging as a neural network mechanism for behavioural tagging. *Nat. Commun.* **7**, 12319 (2016).
47. T. I. Brown, V. A. Carr, K. F. LaRocque, S. E. Favila, A. M. Gordon, B. Bowles, J. N. Bailenson, A. D. Wagner, Prospective representation of navigational goals in the human hippocampus. *Science* **352**, 1323–1326 (2016).
48. R. L. Buckner, The role of the hippocampus in prediction and imagination. *Annu. Rev. Psychol.* **61**, 27–48 (2010).
49. H. F. Ólafsdóttir, C. Barry, A. B. Saleem, D. Hassabis, H. J. Spiers, Hippocampal place cells construct reward related sequences through unexplored space. *Elife* **4**, e06063 (2015).
50. R. A. Diana, A. P. Yonelinas, C. Ranganath, Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends Cogn. Sci.* **11**, 379–386 (2007).
51. M. Osta, D. Lozano-Soldevilla, A. Gil-Nagel, R. Toledano, C. R. Oehrn, L. Kunz, M. Yebra, C. Mendez-Bertolo, L. Stieglitz, J. Sarnthein, N. Axmacher, S. Moratti, B. A. Strange, Aversive memory formation in humans involves an amygdala-hippocampus phase code. *Nat. Commun.* **13**, 6403 (2022).
52. B. M. Asile, V. D. Costa, J. L. Schafroth, C. L. Karaskiewicz, D. R. Lucas, E. A. Murray, The amygdala is not necessary for the familiarity aspect of recognition memory. *Nat. Commun.* **14**, 8109 (2023).
53. T. F. Brady, M. M. Robinson, J. R. Williams, J. T. Wixted, Measuring memory is harder than you think: How to avoid problematic measurement practices in memory research. *Psychon. Bull. Rev.* **30**, 421–449 (2023).
54. A. P. Yonelinas, Consciousness, control, and confidence: The 3 Cs of recognition memory. *J. Exp. Psychol. Gen.* **130**, 361–379 (2001).

Acknowledgments: We express gratitude to C. T. Gill for assistance with the analysis, (Tiffany) Y. Bai for helpful and constructive discussions on writing the manuscript, and the authors of the studies included in our analysis who either shared their collected data or made their data openly accessible. **Funding:** R.M.G.R. is supported by grants from the National Institutes of Health (R01-MH114877, R01-AG063775, and R01-AG082645), the International Obsessive-Compulsive Disorder Foundation (IOCDF), the AE Research Foundation, and philanthropy. The funders had no role in the conceptualization, preparation, or decision to publish this manuscript. **Author contributions:** Conceptualization: C.L., W.W., P.C., and R.M.G.R. Methodology: C.L., W.W., P.C., S.G., and R.M.G.R. Software: C.L. and P.C. Validation: C.L. and W.W. Formal analysis: C.L. and W.W. Investigation: C.L., W.W., P.C., and S.S. Resources: R.M.G.R. Data curation: C.L. and S.S. Writing—original draft: C.L. and W.W. Writing—review and editing: C.L., W.W., P.C., S.G., S.S., and R.M.G.R. Visualization: C.L. and W.W. Supervision: C.L., W.W., and R.M.G.R. Project administration: C.L., S.S., W.W., and R.M.G.R. Funding acquisition: R.M.G.R. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All deidentified data, together with the scripts for fitting linear mixed-effects models and extracting image features via a ResNet-18 CNN, have been deposited in Zenodo and are publicly available at https://doi.org/10.5281/zenodo.15889161.

Submitted 11 April 2025
Accepted 25 August 2025
Published 24 September 2025
10.1126/sciadv.ady1704