Effects of surprising rewards on pattern separation

Dale Zhou (dale.zhou@uci.edu)[1,2,3], **Gimarie Irizarry-Martínez (girizarr@uci.edu)**[2,3],
Shuheng Guo (jianleg@uci.edu)[1], **Soyun Kim (soyunk2@uci.edu)**[2,3], **John Janecek (janecekj@uci.edu)**[2],
Nandita Tuteja (tutejan@uci.edu)[2,3], **Novelle Meza (njmeza@uci.edu)**[2,3], **Liv McMillan (liv@uci.edu)**[2,3],
Julian Thayer (jfthayer@uci.edu)[3,4], **Aaron Bornstein (aaron.bornstein@uci.edu)**[1,3,*] **&**
Michael A. Yassa (myassa@uci.edu)[2,3,4,*]
[1]Department of Cognitive Sciences, University of California, Irvine
[2]Department of Neurobiology and Behavior, University of California, Irvine
[3]Center for the Neurobiology of Learning and Memory, University of California, Irvine
[4]Department of Psychological Science, University of California, Irvine

Abstract

Surprising feedback alters memory of events. Reward prediction errors (RPEs) signal when feedback violates expectations. Although RPEs are known to influence memory, their role in pattern separation—a fundamental computation supporting discriminated episodic representations—remains unclear. We ask whether surprising outcomes strengthen memory encoding and if welcome (positive RPE) versus unwelcome (negative RPE) surprises differentially affect memory discrimination. We designed a variant of an established task developed to probe pattern separation, the Mnemonic Discrimination Task, that introduces post-trial reward and punishment feedback during encoding. We find that while surprising rewards enhance memory discrimination overall, trials with positive RPEs lead to better discrimination of similar lure items, but trials with negative RPEs lead to better recognition of repeated target items. Generalizing expectations to semantically similar stimuli further benefits discrimination. These findings suggest that surprising feedback enhances memory discrimination, with effects depending on the type of feedback and similarity structure of experiences.

Keywords: memory discrimination; reward-modulated memory; reward prediction error; value generalization

Introduction

The ability to discriminate between similar experiences is a key memory and decision-making function (Botvinick et al., 2015; Noh et al., 2014, 2023; Yassa & Stark, 2011). For example, a forager encounters a mushroom that is similar to one that it has eaten safely before. Therefore, the forager eats this mushroom. The forager may find that the mushroom is unexpectedly even more delicious than the one it had seen before, making it more likely to seek out this family of mushrooms in the future. Alternatively, eating this mushroom can make the forager unexpectedly sick. Do surprising outcomes like this make it more likely that the forager will remember the new mushroom and be able to discriminate it from other similar mushrooms in the future? The answer to this question may depend not just on the magnitude of the surprising outcome, but also its sign and how similar the stimulus is to future experiences.

Reward expectations influence memory encoding. Events with a surprisingly high or low reward can be quantified as having a high reward prediction error (RPEs) from reinforcement learning models (Rescorla, 1972; Schultz, 1998). Using such models of reward expectations have been shown to be linked to memory recognition, recall, and event segmentation (Rosenbaum et al., 2022; Rouhani et al., 2018, 2020). However, it is less understood whether and how RPEs affect the discrimination of similar memories. Indeed, previous reports differ on whether surprising rewards influence memory primarily when their sign is positive (Calderon et al., 2021; De Loof et al., 2018; Jang et al., 2019), or regardless of sign (Chen et al., 2025; Rouhani et al., 2020).

Intuitively, the similarity between previous and current sensory experiences, as well as their sign, should both influence how and what we remember. Prior theoretical and empirical work has suggested that learned reward values sometimes *generalize* from the specific experience to novel, related encounters (Wu et al., 2025); while other work has identified situations where welcome or unwelcome surprising outcomes lead to more detailed memory for the specific event (Pupillo & Bruckner, 2023). Here, we exploited the randomization of stimulus and reward sequences to study how varied similarity between past and present stimuli as well as the sign and magnitude of surprising feedback contribute distinctly to subsequent memory discrimination and recognition performance.

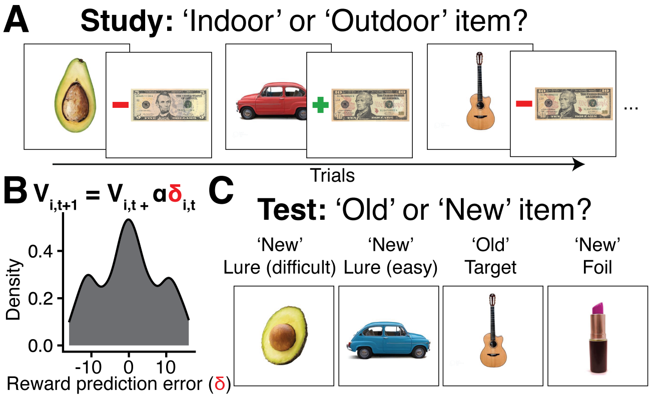

Figure 1: **Task schematic. (A)** Exposure to objects followed by positive or negative reward feedback. **(B)** Distribution of reward prediction errors, quantifying how surprising a reward is per trial, across trials and participants. **(C)** Trial types and their correct responses.

To accomplish this goal and attempt to resolve seeming discrepancies in previous reports, we developed a novel mnemonic discrimination task (MDT) that can be used to assess modulation of memory by stimulus value. MDTs probe behavioral signatures of pattern separation ability between objects that are perceptually similar ("lures") to a previously encoded item in memory ("first presentation") (Leal & Yassa, 2018). Participants are exposed to first presentations of every day items (standard for all MDTs) during a study phase. Then during a testing phase, they see either similar "lures", repeated "targets," or brand new "foils" **(Figure 1)**. Discrimination performance, measured by the proportion of correctly identified lures corrected for a response bias for 'new' responses (('new'|lure)-('new'|target)), is thought to assess detail knowledge or specific recollection (Norman, 2010; Yonelinas et al., 2010). We designed a variant, the *value-modulated pattern separation task*, to investigate the potential effects of reward and reward-related computations during encoding on subsequent memory. This task variant adds a randomized reward that follows each study-phase image. We hypothesized that surprising rewards would enhance later lure discrimination and target recognition because surprising rewards are known to enhance memory encoding (Murty & Adcock, 2014; Murty et al., 2016; Rouhani et al., 2018) . This unexpected reward signal could strengthen the trace for the original stimulus, which could reduce its vulnerability to interference from a similar trace during the test phase. We were further interested in testing whether the experience of reward would generalize to perceptually and semantically-related images (Wu et al., 2025; Zhou et al., 2025), and whether this generalization would interact with the sign of the surprising rewards (Jang et al., 2019; Rouhani & Niv, 2021).

Methods

Value-modulated pattern separation task

The task included a study phase (135 trials) followed by a testing phase (180 trials). During the study phase, participants saw one object per trial. As a cover task during learning, they judged whether each object is typically found indoors or outdoors. Participants were informed that artificial intelligence determines the reward based on whether their judgment was correct or incorrect in responding indoor or outdoor.

In actuality, the reward sequence was determined according to a stationary (uniform) probability distribution over four reward outcomes: -10, -5, 5 or 10. This simple reward distribution was used to simulate possible reward sequences, which were then filtered to the sequences that have the smallest fluctuations in the absolute expected value as well as the lowest correlation between outcomes and expected values (detailed in the *Rescorla-Wagner reward learning* section below). This allowed us to control prediction errors to be more consistent across participants while retaining variation in the prediction errors of lure trials for analyses of individual differences. This design also better enables us to model responses to the magnitude of reward outcomes more independently from responses to learned expectations (e.g. reducing colinearity of covariates in linear models).

During the testing phase, participants viewed an object per trial and indicated whether they have not seen it ("new") or if they have previously seen it ("old") during the study phase. Crucially, some of the objects they viewed were foils, which were completely novel, some were repeated targets that they previously viewed during learning, and some were lures, which had varying degrees of similarity (easy and difficult) to a prior object that they viewed during the study phase. Easy lures are more "mnemonically dissimilar" than difficult lures as previously determined by binned lure discrimination performance in an independent sample (Lacy et al., 2011).

The primary performance measure of memory discrimination is the *Lure Discrimination Index* (LDI), or the proportion of correct rejections adjusted by a response bias: $p(\text{"new"}|\text{lure}) - p(\text{"new"}|\text{target})$. Target recognition memory is the bias-corrected hit rate: $p(\text{"old"}|\text{target}) - p(\text{"old"}|\text{foil})$.

Participants

An online sample from Prolific (n=125; average age=34 ± 12 years old; 73 female, 52 male) completed the task. Participants received monetary compensation for participating. Participants were required to complete a brief tutorial. Immediately in between the learning and testing phases, they had to respond to two practice trials until both were correct before proceeding. Participant exclusion criteria included major medical conditions, neurological disorders and head injury (n=7). We also excluded participants based on responding to fewer than 80% of trials (n=5) and having negative LDI (lure correct rejection worse than the response bias for 'new' on target trials) (n=15). These steps led to our final sample (n=98; 34 ± 11 years old; 56 female, 42 male).

Models

Rescorla-Wagner reward learning As mentioned above, prediction errors were calculated on a per-trial basis from a random reward sequence. During the study phase, we implemented a Rescorla-Wagner reinforcement learning model (Rescorla, 1972). On each trial t, the model updates its expected value V for the presented image i according to the following update rule:

$$V_{i,t+1} = V_{i,t} + \alpha\, \delta_{i,t}, \tag{1}$$

where $V_{i,t}$ denotes the predicted value of image i prior to trial t and $\delta_{i,t}$ is the reward prediction error (RPE) for image i and trial t, defined as the difference between the received outcome $R_{i,t}$ and the expected value:

$$\delta_{i,t} = R_{i,t} - V_{i,t}. \tag{2}$$

Positive δ indicates that the received outcome was greater than expected, whereas a negative δ indicates that the received outcome was less than expected. In addition to positive or negative δs, we further analyze unsigned prediction errors because their magnitude has been shown to index memory and processing speed (Liu et al., 2025; Rosenbaum et al., 2022;

Rouhani et al., 2018). We constructed reward sequences such that trials had balanced δ values (mean $\delta = 0$) and minimal $V_{i,t}$ to minimize the effect of large variations or skewed distributions across trials.

Lastly, α is the learning rate ranging from 0 to 1. As the task by design did not make reward contingent on the particular choice, we simulated a range of possible α to determine how learning rates modulate relationships between RPE and memory. In stationary environments, including our uniform distribution of rewards, the optimal α is expected to be low, such that learning updates occur gradually rather than adjusting strongly to random fluctuations in the reward sequence (Simoens et al., 2024). Further, we also examined whether the relationship between RPE and memory performance differed under a set of plausible learning rates, as humans maintain multiple representations of the environment that differ in how much past experience they rely on, each generating its own prediction errors (Bornstein & Daw, 2012, 2013; Diuk et al., 2013; Gershman et al., 2009; Gläscher & Büchel, 2005; Wilson et al., 2013).

For analyses of the testing phase performance, we mapped the RPE from each study phase trial's first presentation to its matched test phase stimulus (either lure or repeated target). Foils did not have an associated RPE for an original stimulus during the study phase so were not analyzed.

Simulating the effects of different fixed learning rates Given the random reward sequence observed by each participant, we determine the optimal α from simulations using a reinforcement learning model with five optimization functions.

(1) Learn the mean of the underlying reward distribution:

$$\alpha_1^* = \arg\min_{\alpha} \left[\frac{1}{T} \sum_{t=1}^{T} (V_t - \mu)^2 \right] \text{ where } \mu = \frac{1}{T} \sum_{t=1}^{T} R_t,$$

(2) predict the next reward:

$$\alpha_2^* = \arg\min_{\alpha} \left[\frac{1}{T-1} \sum_{t=1}^{T-1} (V_t - R_{t+1})^2 \right],$$

(3) minimize overall surprise:

$$\alpha_3^* = \arg\min_{\alpha} \left[\frac{1}{T} \sum_{t=1}^{T} (V_t - R_t)^2 \right],$$

(4) maximize overall expected value:

$$\alpha^* = \arg\max_{\alpha} \sum_{t=1}^{T} V_t$$

and (5) prioritize recent memories:

$$\alpha_4^* = \arg\min_{\alpha} \left[\frac{\sum_{t=1}^{T-1} \gamma^{T-t} (V_t - R_{t+1})^2}{\sum_{t=1}^{T-1} \gamma^{T-t}} \right] \text{ where } \gamma = 0.95.$$

We used these simulations to interpret the computational usage of a range of learning rates.

Similarity-based value generalization Reward expectations are driven both by the history of rewards as well as the similarity of the current observation to past observations. Therefore, we expanded the base Rescorla-Wagner model to also generate RPEs from value updates not only to the currently observed stimulus but also to all similar stimuli.

Similarity-based generalization models make quantitative predictions about how much the value of one stimulus should be generalized to another based on their feature similarity (Wu et al., 2025). To determine feature-based similarity across stimuli, we used convolutional neural networks (CNNs) trained with supervised learning to classify images that are known to form interpretable representations of perceptual (e.g. edges and textures of the silhouette of a cat) and semantic features (e.g. parts and configurations of the eyes, nose, and whisker of a cat) (Krizhevsky et al., 2012). We considered detailed perceptual features (first layer activations) extracted from a pre-trained deep CNN. We also considered semantic features from a pre-trained vision-text transformer using the contrastive language-image pre-training (CLIP) neural network framework with the ViT-H/14 image embedding variant (Cherti et al., 2023).

For each image, a fixed-dimensional embedding vector was extracted and normalized to unit length. Stimulus similarity was quantified using cosine similarity between normalized embedding vectors. The similarity kernel K was defined as:

$$K_{ij} = \mathbf{e}_i^{\top} \mathbf{e}_j, \tag{3}$$

where $\mathbf{e}_i$ is the embedding of stimulus i, normalized such that $\|\mathbf{e}_i\| = 1$, and K is a symmetric stimulus-by-stimulus similarity matrix with a diagonal of 1.

After observing stimulus i_t, values were updated to other stimuli j according to:

$$V_j \leftarrow V_j + \gamma \, \mathrm{PE}_t \, K_{i_t j}, \tag{4}$$

where j indexes all stimuli, $K_{i_t j}$ is the cosine similarity between the current stimulus and stimulus j, and γ is a parameter controlling the magnitude of similarity-based updating (set to 0.1 for modest 10% generalization of current stimulus to similar stimuli). The updating for the experienced stimulus i is updated as in the base Rescorla-Wagner model ($K_{i_t i_t} = 1$).

Congruency of indoor/outdoor context Finally, we considered an additional model to capture potential confounds introduced by the reward feedback on the indoor/outdoor cover task. The confound concerns study phase trials in which participants are punished for judging an unambiguously "indoor" object like a refrigerator as an indoor object. Assuming participants believe their own judgment is correct, a similar effect is not expected for any trials with a positive reward (e.g., endorsing an unambiguously outdoor image as "indoor" and receiving positive feedback). This incongruency of prior knowledge about the object and the reward outcome may reduce engagement and hinder encoding (Greve et al., 2019; Ortiz-Tudela et al., 2023); or it may represent another form of prediction error and enhance encoding.

Therefore, for each base image i, we obtained a numeric indoor/outdoor rating d_i by entering the name and image of each object into a prompt for an LLM (GPT-4o) to judge the image's typical indoor or outdoor context. "Is this an image of something that's usually found indoors or outdoors? Return your answer in a rating from 1 to 5, where 1 is the most frequently indoor (for example, toaster, dishwasher, or refrigerator), 3 is sometimes found indoors or outdoors (for example, shoes or clocks), and 5 is the most frequently outdoor (for example, a tree or a playground slide)."

Each LLM indoor/outdoor rating was recoded to a –1 (indoor) to +1 (outdoor) scale. Participants' judgments were recoded in the same way. The absolute difference between expectation and choice was calculated as the indoor/outdoor error ζ (range [0,2]). Finally, reward R was recoded to range [0,1] such that the absolute difference between them treats negative feedback trials with low error as the most surprising with the highest incongruency:

$$\mathrm{PE}_t^{\text{indoor/outdoor}} = \mathrm{PE}_t\left(|\zeta_t - R_t|\right) \tag{5}$$

where PE_t is the standard RPE on trial t, and $|\zeta_t - R_t|$ is largest when most heavily punished (reward $= -10$) despite judging an object to be indoor or outdoor in a manner congruent with automated ratings (error $= 0$).

Results

Memory performance

Individuals varied memory discrimination performance on lures and target recognition. The mean LDI was 0.15 ± 0.09 **(Figure 2A)**. As expected, performance on difficult stimuli was worse than on easy stimuli ($t(97) = -3.7, p = 0.0004$). The mean target recognition was 0.35 ± 0.11 **(Figure 2B)**. We did not observe a correlation between the trial-by-trial reward and lure correct rejections ($b = 0.02, p = 0.30$) nor reward and target recognition ($b = -0.03, p = 0.42$; **Figure 2C-D**).

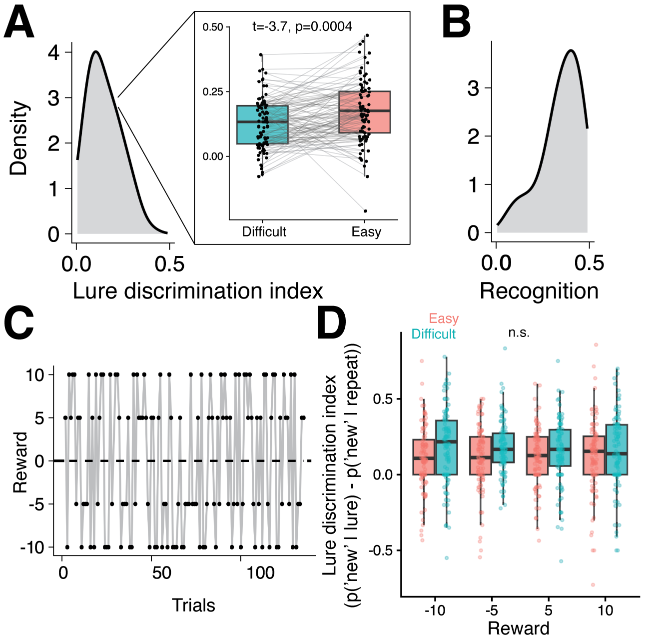

Figure 2: **Task performance.** **(A)** Individual differences in memory discrimination performance. Performance was worse for discriminating more mnemonically similar objects ("difficult") versus more mnemonically dissimilar objects ("easy"). Lines connect data points which represent the same individual. **(B)** Individual differences in target recognition. **(C)** An example reward sequence. **(D)** LDIs recalculated over the subset of trials that obtained each reward value. No relationship between reward and lure discrimination indices, nor when splitting by trial difficulty.

Simulated learning rates for value learning

To go beyond evaluating raw reward values, reward expectations can be estimated from Rescorla-Wagner models using a fixed learning rate across trials. We sought to simulate several possibilities of learning rates that could optimize for different learning objectives. Each participant's reward sequence was used as the input for computing a cost based on objective functions related to learning, prediction, surprise, value, and memory targets **(Figure 3A)**. In short, these objectives consider that participants may process reward signals to accurately learn the average reward, predict the next trial's reward, minimize overall surprise, maximize expected value, or prioritize the outcomes of recent memories. With our random, stationary probability distribution of rewards, it is theoretically optimal to form reward expectations with a low (slow) learning rate, because they make adjustments to reward expectations less sensitive to the random variations in our reward sequence. Consistent with this idea, a grid search using a step size of 0.001 returned relatively low learning rates, with a median $\alpha = 0.03$ (prioritizing recent memories and maximizing expected value) and 0.11 (predicting the next reward, minimizing error, and estimating the mean) across objectives.

Surprising rewards enhance memory discrimination

We used these two identified learning rates to interpret the computational usage of RPEs generated across a range of learning rates **(Figure 3B)**. We tracked the partial correlation of RPEs and memory performance, including a covariate controlling for age because age has known effects on these memory measures. At $\alpha = 0.03$, RPEs were significantly associated with LDI ($\beta = 0.24, p = 0.04$), marginally with target recognition ($\beta = 0.24, p = 0.08$), and not associated with reaction time ($\beta = 0.02, p = 0.91$). When RPEs were generated using $\alpha = 0.11$, the relationship between RPE and reaction time was strongest though marginal ($\beta = 0.24, p = 0.14$). At this α, there was also a significant relationship between RPE and LDI ($\beta = 0.17, p = 0.004$; **Figure 4B**) as well as target recognition ($\beta = 0.19, p = 0.02$). Lastly, we considered the possibility that participants completely based their value estimate based on the outcome of the preceding trial, such that $\alpha = 1$. Under this assumption, there was no relationship with LDI ($\beta = 0.02, p = 0.17$), whereas RPEs had a positive re-

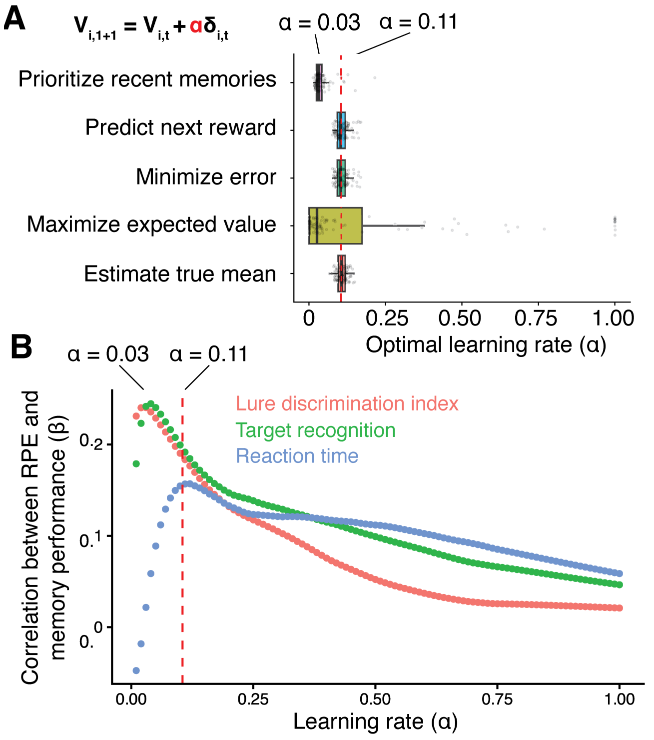

Figure 3: **Simulating learning rates for different reward learning objectives. (A)** Each participant's observed reward sequence was used to calculate the cost associated with different objective functions based on the Rescorla-Wagner model, identifying median $\alpha = 0.03$ and 0.11. The maximization of expected value depends on how positive or negative rewards are front or back-loaded in the reward sequence, resulting in greater variance across participants. **(B)** The strongest relationships between RPEs and memory performance are observed when using the identified learning rates.

lationship with target recognition ($\beta = 0.05, p = 0.011$) and with reaction time ($\beta = 0.06, p = 0.016$). These simulations suggest a range of computationally relevant learning rates for memory processes.

Positive RPEs enhance memory discrimination Given that the amount of reward was not related to LDIs on either easy or difficult trials, we next assessed whether the surprising or unexpectedly positive or negative rewards affect memory on a trial-by-trial basis. There was a trial-by-trial relationship between signed RPE and lure correct rejections ($b = 0.05, p = 0.046$; **Figure 4C**), suggesting that trials with more positive RPEs had enhanced memory discrimination. This effect was robust to differences in learning rates (range of $\beta = [0.04, 0.05]$) with a decrease in effect size for RPEs generated with larger learning rates ($r = -0.75$). Surprising rewards (unsigned RPEs) were negatively related to lure correct rejections ($b = -0.05, p = 0.03$), suggesting that inaccurate reward expectations for a trial can interfere with memory discrimination even while more surprising rewards overall appear to improve it across participants.

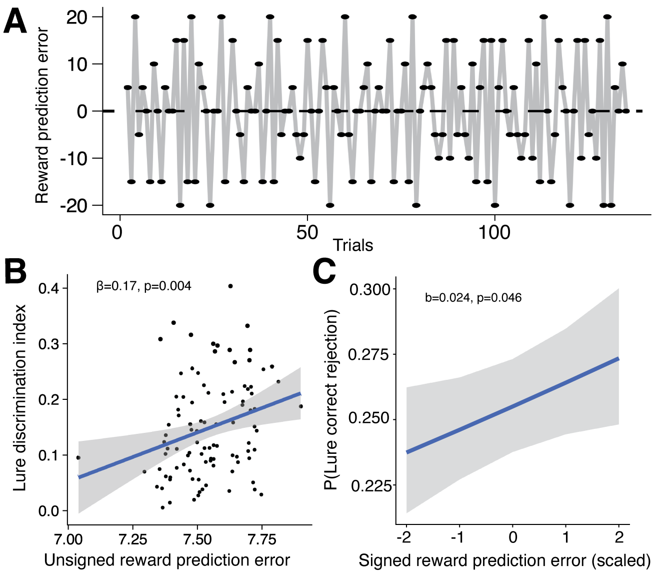

Figure 4: **Surprising rewards enhance memory discrimination. (A)** An example RPE sequence. **(B)** Larger RPEs magnitudes overall were associated with better memory discrimation across participants. **(C)** More positive RPEs were associated with a greater probability of lure correct rejection across lure trials.

Negative RPEs enhance target recognition Unwelcome surprises (negative RPEs) were associated with better target recognition. There was a trial-by-trial relationship between signed RPE and target hits ($b = -0.12, p = 0.006$), but not unsigned RPE and repeat hits ($b = 0.03, p = 0.43$). The effect of signed RPE was robust to differences in learning rates and consistently negative (range of $\beta = [-0.15, -0.10]$) with a decrease in effect size with greater learning rates ($r = -0.81$).

Taken together, positive RPE was correlated with enhanced memory discrimination, negative RPE was correlated with enhanced target recognition, and the effects were strongest with slow learning rates which are theoretically optimal for these reward sequences.

Cover task violations There was a greater incongruence was associated with worse target recognition ($\beta = -0.23, p = 0.03$), indicating that violations of real-world context interfered with successful memory encoding perhaps as a distraction. We did not observe an effect on LDI or reaction time, nor did the incongruency value confound the observed relationships between RPEs and lure discrimination when included as an additional covariate.

Reward generalization across similar perceptual and semantic features

Finally, we tested how reward generalization could produce RPEs that differently relate to memory discrimination. Perceptual vs. semantic feature generalization interacted with

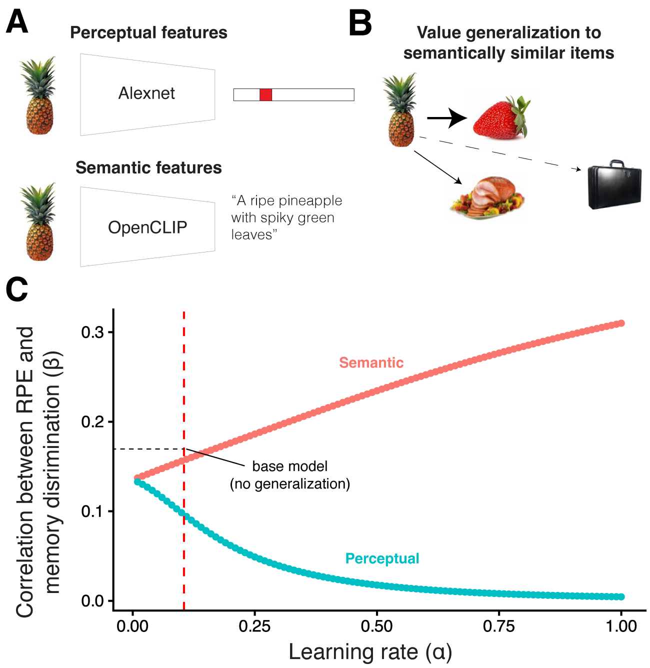

Figure 5: **Differential effects of value generalization across perceptual or semantic features depending on learning rate.** **(A)** Perceptual and semantic features were extracted from images using pre-trained neural networks. **(B)** Value generalizes to similar stimuli according to perceptual or semantic feature similarity. **(C)** Perceptual feature generalization appears to benefit memory discrimination most at lower learning rates whereas semantic feature generalization benefits memory discrimination most at higher learning rates. Red vertical dashed line indicates $\alpha = 0.11$ identified from the prior simulations; black horizontal dashed line indicates the effect with no generalization for comparison.

learning rate to vary how RPE modulates memory discrimination **(Figure 5)**. With perceptual feature generalization, the relationship between RPE and memory discrimination was strongest with lower learning rates. In contrast, with semantic feature generalization, the relationship between RPE and memory discrimination was strongest with higher learning rates. A higher learning rate may benefit memory discrimination by quickly generalizing expectations across stimuli with similar semantic features.

Discussion

We found evidence that surprising rewards enhance both the discrimination and recognition of memory encoding across individuals. Better lure discrimination and target recognition was associated with stronger surprise. Surprising events that violate predictions were also associated with longer reaction times, which may involve additional belief updating and processing for response selection. We identified a range of computationally relevant learning rates for learning from the task's random reward sequence. The strongest correlations between RPE and memory measures converged on the RPEs produced by the identified learning rates.

These findings are consistent with other reports that surprising reward events may enhance memory (Murty & Adcock, 2014; Murty et al., 2016). Similar to prior work, we found effects of unsigned RPE (Rouhani et al., 2018) and signed RPE (Calderon et al., 2021; De Loof et al., 2018; Jang et al., 2019) on memory measures. We did not find a main effect of continuous reward values, or reward versus punishment, on memory performance. In contrast, a different task used lures to elicit surprise and test recognition memory of lures, showing that rewards enhance recognition of the surprising lures more than punishments (Murty et al., 2016). Rather than using lures to elicit surprise, our findings highlight how surprising violations of reward expectations, such as better-than-expected rewards, enhance memory discrimination for lures. Generalizing expectations to future situations with similar semantic features may further support memory discrimination, consistent with prior work suggesting that stimulus features can guide value generalization and augment episodic memory (Greve et al., 2019; Wu et al., 2025)

In future work, neuroimaging will be used with this task to study circuits where pattern separation and reward prediction error processes interact. Relevant regions of interest will likely involve the ventral tegmental area releasing dopaminergic signals of positive and negative RPE (Montague et al., 1996), the locus coeruleus releasing dopaminergic and noradrenergic signals of surprise magnitude (Clewett et al., 2018; Jordan, 2024; Jordan & Keller, 2023), and the hippocampus separating similar and interfering memory traces (Chen et al., 2025; Jordan, 2024; Jordan & Keller, 2023; Lisman & Grace, 2005; Murty & Adcock, 2014; Nassar et al., 2012).

This study had several strengths and limitations. Strengths included tightly controlled mnemonic similarity between lures and target stimuli as well as randomized reward sequences that link reward processing to a canonical memory discrimination task (Lacy et al., 2011). Limitations include data from an online sample that is highly heterogeneous and may underestimate true task engagement. Correct responses were overall lower relative to traditional MDT designs that are used in the lab (Stark et al., 2023). Moreover, several results depend on specific fixed learning rates. We found some evidence that the "correctness" of indoor/outdoor judgments may introduce a distraction that interferes with memory encoding, consistent with work showing effects of congruent vs. incongruent prior knowledge on memory (Greve et al., 2019). It is possible that the four discrete reward outcomes did not appear meaningfully tied to actions for the participants, which may blunt the association of rewards or reward expectations with the preceding object. Future designs could address these limitations by using a continuous rather than discrete reward distribution linked to independent raters' indoor/outdoor judgments while randomizing the reward amount separately for congruent versus incongruent judgments.

References

Bornstein, A. M., & Daw, N. D. (2012). Dissociating hippocampal and striatal contributions to sequential prediction learning. *European Journal of Neuroscience*, *35*(7), 1011–1023.

Bornstein, A. M., & Daw, N. D. (2013). Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. *PLoS computational biology*, *9*(12), e1003387.

Botvinick, M., Weinstein, A., Solway, A., & Barto, A. (2015). Reinforcement learning, efficient coding, and the statistics of natural tasks. *Current opinion in behavioral sciences*, *5*, 71–77.

Calderon, C. B., De Loof, E., Ergo, K., Snoeck, A., Boehler, C. N., & Verguts, T. (2021). Signed reward prediction errors in the ventral striatum drive episodic memory. *Journal of Neuroscience*, *41*(8), 1716–1726.

Chen, Y., Zhao, C., Liu, R., & Li, Q. (2025). Unexpected feedback enhances episodic memory: Exploring signed and unsigned reward prediction errors with eeg. *NeuroImage*, 121303.

Cherti, M., Beaumont, R., Wightman, R., Wortsman, M., Ilharco, G., Gordon, C., Schuhmann, C., Schmidt, L., & Jitsev, J. (2023). Reproducible scaling laws for contrastive language-image learning. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2818–2829.

Clewett, D. V., Huang, R., Velasco, R., Lee, T.-H., & Mather, M. (2018). Locus coeruleus activity strengthens prioritized memories under arousal. *Journal of Neuroscience*, *38*(6), 1558–1574.

De Loof, E., Ergo, K., Naert, L., Janssens, C., Talsma, D., Van Opstal, F., & Verguts, T. (2018). Signed reward prediction errors drive declarative learning. *PLoS One*, *13*(1), e0189212.

Diuk, C., Tsai, K., Wallis, J., Botvinick, M., & Niv, Y. (2013). Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. *Journal of Neuroscience*, *33*(13), 5797–5805.

Gershman, S. J., Pesaran, B., & Daw, N. D. (2009). Human reinforcement learning subdivides structured action spaces by learning effector-specific values. *Journal of Neuroscience*, *29*(43), 13524–13531.

Gläscher, J., & Büchel, C. (2005). Formal learning theory dissociates brain regions with different temporal integration. *Neuron*, *47*(2), 295–306.

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General*, *148*(2), 325.

Jang, A. I., Nassar, M. R., Dillon, D. G., & Frank, M. J. (2019). Positive reward prediction errors during decision-making strengthen memory encoding. *Nature human behaviour*, *3*(7), 719–732.

Jordan, R. (2024). The locus coeruleus as a global model failure system. *Trends in Neurosciences*, *47*(2), 92–105.

Jordan, R., & Keller, G. B. (2023). The locus coeruleus broadcasts prediction errors across the cortex to promote sensorimotor plasticity. *Elife*, *12*, RP85111.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.

Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. (2011). Distinct pattern separation related transfer functions in human ca3/dentate and ca1 revealed using high-resolution fmri and variable mnemonic similarity. *Learning & memory*, *18*(1), 15–18.

Leal, S. L., & Yassa, M. A. (2018). Integrating new findings and examining clinical applications of pattern separation. *Nature neuroscience*, *21*(2), 163–173.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-vta loop: Controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703–713.

Liu, F., Jiang, Y., & Du, B. (2025). Reward prediction-error promotes the neural encoding of episodic learning. *Neuropsychologia*, *211*, 109120.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of neuroscience*, *16*(5), 1936–1947.

Murty, V. P., & Adcock, R. A. (2014). Enriched encoding: Reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cerebral Cortex*, *24*(8), 2160–2168.

Murty, V. P., LaBar, K. S., & Adcock, R. A. (2016). Distinct medial temporal networks encode surprise during motivation by reward versus punishment. *Neurobiology of learning and memory*, *134*, 55–64.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature neuroscience*, *15*(7), 1040–1046.

Noh, S. M., Singla, U. K., Bennett, I. J., & Bornstein, A. M. (2023). Memory precision and age differentially predict the use of decision-making strategies across the lifespan. *Scientific Reports*, *13*(1), 17014.

Noh, S. M., Yan, V. X., Vendetti, M. S., Castel, A. D., & Bjork, R. A. (2014). Multilevel induction of categories: Venomous snakes hijack the learning of lower category levels. *Psychological science*, *25*(8), 1592–1599.

Norman, K. A. (2010). How hippocampus and cortex contribute to recognition memory: Revisiting the complementary learning systems model. *Hippocampus*, *20*(11), 1217–1227.

Ortiz-Tudela, J., Nolden, S., Pupillo, F., Ehrlich, I., Schommartz, I., Turan, G., & Shing, Y. L. (2023). Not what u expect: Effects of prediction errors on item memory. *Journal of Experimental Psychology: General*, *152*(8), 2160.

Pupillo, F., & Bruckner, R. (2023). Signed and unsigned effects of prediction error on memory: Is it a matter of choice? *Neuroscience & biobehavioral reviews*, *153*, 105371.

Rescorla, R. A. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning, Current research and theory*, *2*, 64–69.

Rosenbaum, G. M., Grassie, H. L., & Hartley, C. A. (2022). Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. *ELife*, *11*, e64620.

Rouhani, N., & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory. *elife*, *10*, e61077.

Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1430.

Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020). Reward prediction errors create event boundaries in memory. *Cognition*, *203*, 104269.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*.

Simoens, J., Verguts, T., & Braem, S. (2024). Learning environment-specific learning rates. *PLoS computational biology*, *20*(3), e1011978.

Stark, C. E., Noche, J. A., Ebersberger, J. R., Mayer, L., & Stark, S. M. (2023). Optimizing the mnemonic similarity task for efficient, widespread use. *Frontiers in behavioral neuroscience*, *17*, 1080366.

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2013). A mixture of delta-rules approximation to bayesian inference in change-point problems. *PLoS computational biology*, *9*(7), e1003150.

Wu, C. M., Meder, B., & Schulz, E. (2025). Unifying principles of generalization: Past, present, and future. *Annual Review of Psychology*, *76*(1), 275–302.

Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in neurosciences*, *34*(10), 515–525.

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

Zhou, D., Noh, S. M., Harhen, N. C., Banavar, N. V., Kirwan, C. B., Yassa, M. A., & Bornstein, A. M. (2025). A compressed code for memory discrimination. *arXiv preprint arXiv:2510.10791*.