Psychological Review

Cognitive Constraints and Reward Environments Jointly Shape Memory Formation

Si Ma, Vencislav Popov, and Qiong Zhang

Online First Publication, March 19, 2026. https://dx.doi.org/10.1037/rev0000612

CITATION

Ma, S., Popov, V., & Zhang, Q. (2026). Cognitive constraints and reward environments jointly shape memory formation. *Psychological Review*. Advance online publication. https://dx.doi.org/10.1037/rev0000612

AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychological Review

ISSN: 0033-295X

https://doi.org/10.1037/rev0000612

THEORETICAL NOTE

Cognitive Constraints and Reward Environments Jointly Shape Memory Formation

Si Ma[1], Vencislav Popov[2], and Qiong Zhang[1, 3]
[1] Department of Psychology, Rutgers University–New Brunswick
[2] Department of Psychology, University of Zürich
[3] Department of Computer Science, Rutgers University–New Brunswick

A key adaptive feature of human episodic memory is its ability to selectively encode important information. While extensive empirical evidence supports that reward enhances memory, this monotonic relationship does not always hold, and the influence of reward on memory formation can exhibit complex patterns. To reconcile these findings, we present a computational model that considers not only how people maximize their rewards but also the cognitive constraints underlying this adaptation. Unlike previous theoretical accounts that assume a direct link between reward value and memory encoding strength, our model adaptively decides how strongly to encode each item based on the overall reward environment and its interaction with limited cognitive resources at encoding. We validated the model's predictions across three experiments, successfully explaining why high-reward items do not always have a memory advantage, how reward context influences memory, and the role of reward anticipation. Importantly, the model predictions are parameter-free. They are derived solely from optimal adaptation to a given reward environment, under cognitive constraints independently characterized by previous studies, rather than being fit to the empirical data that the model seeks to explain. These findings support that memory encoding is an active process involving metacognitive control, where limited cognitive resources are strategically allocated to maximize overall cumulative rewards, rather than a passive response to the salience of individual items.

Keywords: reward, memory encoding, metacognition, resource-rational analysis

The ability to selectively encode information is an adaptive feature of episodic memory. It is rational to prioritize resources and effort in encoding information that leads to higher rewards. Previous studies found that people better remember information if it will help them gain rewards (Bowen et al., 2020; Gong & Li, 2014; Grandoit et al., 2024; Manga et al., 2020; Middlebrooks et al., 2017; Talmi et al., 2021). Multiple accounts aim to explain this phenomenon of reward-enhanced memory. High-reward items may capture greater attention (Allen & Ueno, 2018; Sandry & Ricker, 2020), be more deeply encoded (Castel, 2007; Cohen et al., 2014), promote a more effective encoding strategy (Hennessee et al., 2019), or receive more selective study (Castel et al., 2013). These accounts are consistent with computational implementations in which memory strength increases when reward increases (Talmi et al., 2021; Zhou et al., 2023). However, this monotonic relationship does not always hold, and the influence of reward on memory formation can exhibit more complex patterns. In this work, we propose a novel account of reward's effect on memory that can reconcile a range of empirical findings.

While high-reward items are generally better remembered than low-reward items, Talmi et al. (2021) discovered an intriguing exception. This memory advantage only exists when high- and low-reward items

---

Zhe-Xin Xu served as action editor.

Si Ma https://orcid.org/0000-0002-0046-7140
Qiong Zhang https://orcid.org/0000-0001-9062-9571

A subset of the results has appeared in a conference article in the non-archival conference proceedings of the Annual Meeting of the Cognitive Science Society in July 2025. Experiment III was preregistered (https://aspredicted.org/jyx6-mdqh.pdf). All data and analysis codes are publicly available on the Open Science Framework at https://osf.io/y6cqt/.

Si Ma played a lead role in data curation, formal analysis, software, and validation and an equal role in conceptualization, investigation, methodology, writing–original draft, and writing–review and editing. Vencislav Popov played a supporting role in conceptualization and writing–review and editing. Qiong Zhang played a lead role in supervision and funding acquisition and an equal role in conceptualization, investigation, methodology, writing–original draft, and writing–review and editing.

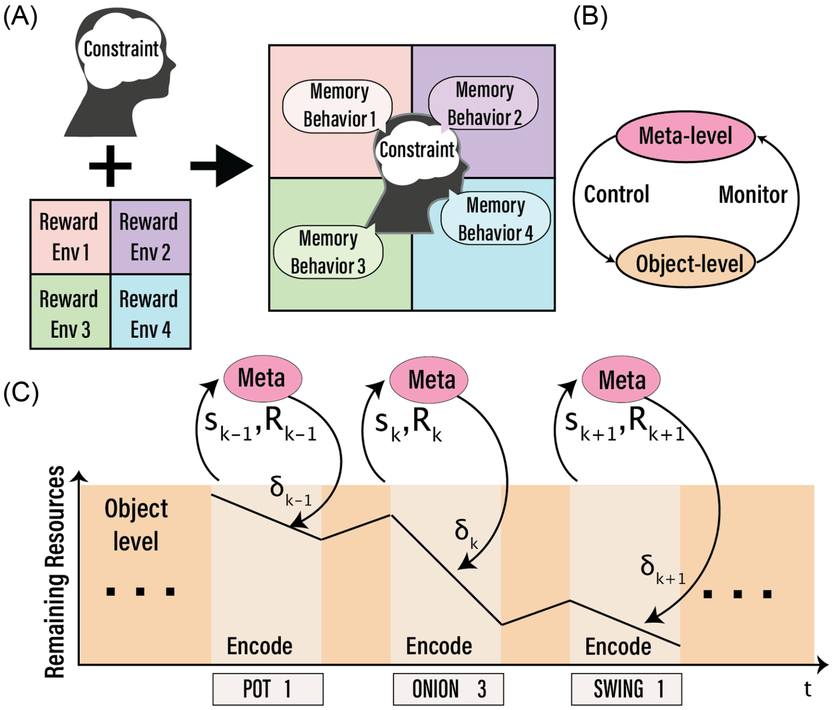
Correspondence concerning this article should be addressed to Qiong Zhang, Department of Psychology, Rutgers University–New Brunswick, 152 Frelinghuysen Road, Piscataway, NJ 08854, United States. Email: qiong.z@rutgers.edu

are mixed within the same list (e.g., "LHLLHLL" where "L" represents a low-reward item and "H" represents a high-reward item; mixed-list condition), but not when they are presented in separate lists (e.g., "LLLLLLL" or "HHHHHHH"; pure-list condition). These effects have been replicated in several follow-up studies (Hellerstedt & Talmi, 2022; Hellerstedt et al., 2023). This raises an important question: If people are rational in selectively remembering more valuable information, why would they recall an equal amount from a high-reward list compared with a low-reward list? People are not only rational and aim to maximize their total rewards in a given task environment; people are also beings with cognitive limits and constraints—as "Human rational behavior is shaped by scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (Simon, 1990, p. 7). We hypothesize that the complex relationships between rewards and memory emerge from people's optimizing total reward gain while operating under cognitive limits and constraints. In other words, if we know people's constraints during memory encoding, and the structure of the reward environment (such as mixed or pure lists in the study by Talmi et al., 2021), we can predict memory behavior in any given reward environment, assuming that people are rational in optimizing their rewards, as illustrated in Figure 1A.

To test this hypothesis, we built a metacognitive model of memory encoding and validated its predictions in three different reward environments (including the difference between pure-list and mixed-list conditions; Talmi et al., 2021). We draw inspiration from early theoretical frameworks in the metamemory literature (Nelson, 1990), which posit that our memory system includes not only an object-level component that describes the detailed processes and constraints of memory but also a meta-level component that adaptively controls ongoing memory processes by monitoring the state of the object level (see Figure 1B). Such a metacognitive model can implement the framework proposed in Figure 1A, as the meta-level component aims to optimize overall task performance (e.g., total reward gain), and the

Figure 1

Our Proposed Theoretical Framework



Note. (A) Adaptive memory behavior is jointly shaped by cognitive constraints during encoding and the reward environment one is in. (B) The interaction between reward adaptation in a given environment and cognitive constraints can be implemented as two interrelated processes of metamemory. The object-level component incorporates information about the cognitive constraints during memory encoding. The meta-level component monitors the state of memory encoding and controls how much to encode each item to optimize overall rewards. (C) The interaction between the object level and the meta level is a continuous process. As the model encounters each item one after another, the meta level determines how much resources to allocate (δ_k) based on the state of memory encoding s_k and the reward information R_k from the object level. The object level then uses the allocated resources (as shown from depletion of cognitive resources at each item encoding) to strengthen the memory. This process repeats until the end of the list, when the model has to recall all the list items. Object-level parameters describe cognitive constraints during memory encoding, which we inherited from a previous resource-based memory model (Ma et al., 2024). We then simulated the optimal allocation of resources across items with different lists and reward structures in each experiment. See the online article for the color version of this figure.

object-level component incorporates information about the cognitive constraints during memory encoding.

What cognitive constraints should be incorporated into the object-level component? While long-term memory storage is vast (Brady et al., 2008), our ability to recall information is constrained by moment-to-moment fluctuations during encoding (Noh et al., 2014; Paller & Wagner, 2002; Sundby et al., 2019). Recent computational models have linked memory performance at recall with the amount of cognitive resources available at encoding (Popov & Reder, 2020; Reder et al., 2000, 2007). The assumption of a finite amount of cognitive resources that need time to recover after depletion (i.e., resource-depletion-and-recovery assumption), as implemented in precise mathematical terms, can account for numerous behavioral (Kowialiewski et al., 2021; Mzrak & Oberauer, 2021; Oberauer, 2022; Popov & Reder, 2020; Popov et al., 2019, 2022) and neural findings (Ma et al., 2024). We hypothesize that the meta-level component continuously monitors the current state of memory encoding from the object level and strategically determines how to allocate the remaining cognitive resources to optimize overall performance.

Critically, when the environment consists of different reward structures, our hypothesis differs from accounts wherein the strength of memory encoding is directly related to the reward magnitude of a given item (Talmi et al., 2021; Zhou et al., 2023). Instead, we posit that the strength of memory encoding is a consequence of optimizing total reward gain through the adaptive allocation of limited cognitive resources. To generate model predictions for specific experiments, we set up the list and reward structure to be identical to the experiment that we aim to model, and ask what the optimal behavior is to encode these items. Importantly, the model has never been directly fitted to the empirical patterns it seeks to explain. Instead, resource constraint parameters are inherited from previous studies (Ma et al., 2024), and the resource allocation behavior is derived through optimization of overall reward, not optimization of data fits. As a result, the model's outcomes serve as genuine predictions.

We validated the model with three experiments. In Experiment I, the model's optimal behavior aligns with observations wherein high-reward items were better remembered than low-reward items in the mixed-list condition, but not in the pure-list condition (Talmi et al., 2021). In Experiment II, we further examined how temporally nearby items within a list compete for the allocation of limited cognitive resources, analyzing a shared data set with reward magnitudes ranging from 1 to 10 (Middlebrooks et al., 2017). We showed that an item's memory performance is not only modulated by its associated reward but also by the rewards of preceding (but not future) items, aligning with the model predictions. In Experiment III, we collected a new data set to test a critical prediction. Although there was no effect of future reward in Experiment II, if people are rational, they should reserve cognitive resources when anticipating high-reward items in the near future. This means that if people can anticipate how much reward future items would have, this should reverse the insensitivity to future rewards, which we demonstrated in Experiment III. Overall, our proposed theoretical framework unifies diverse findings on the role of rewards on memory.

A Metacognitive Model of Memory Encoding

In this section, we specify the details of the proposed metacognitive model. The model aims to capture the encoding processes of a list-learning paradigm (i.e., free recall task), where participants first study a list of items and then attempt to recall items from the list in any order (Murdock, 1962; Roberts, 1972; Standing, 1973). The model consists of two components: an object-level component, responsible for memory encoding, and a meta-level component that monitors and adaptively controls encoding. The interaction between the two components is a continuous process (see Figure 1C). As the model encounters each item one after another, the meta-level determines how much resources to allocate based on the information from the object level, while the object level uses the allocated resources to strengthen the memory trace. This process repeats until the end of the list, when the model has to recall all the list items.

Object-Level Component

At the object level, items are encoded into memory based on the resource-depletion-and-recovery assumption (Ma et al., 2024; Popov & Reder, 2020; Reder et al., 2007). To ensure the cognitive constraints in our study are independently characterized, we adopted the exact model implementation and parameters from a previous study (Ma et al., 2024). The model assumes a finite amount of cognitive resources, which is at its maximum at the start of learning a list of items ($W_{\max} = 1$). Upon studying an item at position k in the list, a fixed proportion (τ) of currently available resources (W_k) is depleted for semantically processing the item, expressed as follows:

$$W_{\text{sem},k} = \tau W_k. \tag{1}$$

Meantime, a fixed proportion (δ) of currently available resources ($W_k - W_{\text{sem},k}$) is allocated to encode the item into episodic memory, expressed as follows:

$$W_{\text{epi},k} = \delta(W_k - W_{\text{sem},k}). \tag{2}$$

The memory strength of the encoded item ($B_{\text{epi},k}$) depends on the amount of resources allocated to the episodic encoding:

$$B_{\text{epi},k} = \sqrt{W_{\text{epi},k}}. \tag{3}$$

The larger the memory strength ($B_{\text{epi},k}$), the higher the probability of recalling the item at position k, expressed as follows:

$$p_k = \Phi\left(\frac{B_{\text{epi},k} - \theta_{\text{epi}}}{\sigma_{\text{epi}}}\right), \tag{4}$$

where Φ represents the cumulative distribution function of the standard normal distribution, with θ_{epi} and σ_{epi} as the mean and the standard deviation used to standardize $B_{\text{epi},k}$.

Once resources are depleted after encoding an item, they need time to recover. Resources recover linearly at a rate of r per second until they reach the maximum $W_{\max}$. Therefore, the amount of resources at the beginning of encoding the next item at position $k + 1$ follows:

$$W_{k+1} = \min(W_{\max}, W_k - W_{\text{sem},k} - W_{\text{epi},k} + rt_k), \tag{5}$$

where t_k is the time between item presentations.

Meta-Level Component

While resource-based memory models traditionally allocate resources passively to each item (e.g., a fixed proportion, δ, of currently available resources; Popov & Reder, 2020; Reder et al., 2000, 2007), the current model allocates resources adaptively through metacognitive control. For each item at position k, the meta level continuously interacts with the object level to determine how much resources to allocate by varying δ (i.e., different δ_k at each position k) to maximize overall reward gain. We obtained the optimal behavior policy in a reinforcement learning framework (Mnih et al., 2015; Sutton & Barto, 2018), where the meta level receives the state information s_k and the reward information R_k from the object level and determines what action δ_k to take (see Figure 1C).

- *States*: In our model, s_k contains the information needed to characterize the current state of memory when encoding an item at position k, including (a) the amount of resources available (i.e., W_k), (b) the property of the current item being encoded (i.e., reward r_k), and (c) how far has it been in the list (i.e., k). The current state additionally contains information about the reward environment specific to each experiment to align the knowledge of the model with that of participants in that experiment.
- *Actions*: δ_k is the action the agent takes when encoding an item at position k. It is the proportion of currently available resources allocated to the item at position k, ranging from 0 to 1.
- *Reward*: The reward received by the agent after taking an action δ_k at item position k is defined as $R_k = p_k \cdot r_k$, where p_k is the item's recall probability (Equation 4) and r_k is its assigned reward value. Although participants in the experiment are only informed of their total reward after the recall test, the model assumes that they can approximate their own recall probabilities by monitoring how well each item was encoded during the study phase.
- *Policy* (π): π is a policy that decides which action to take at a given state s. At each position k, the agent aims to maximize the expected discounted return from the current position k to the end of the list:

$$G_k = \sum_{t=0}^{L-k} \gamma^t R_{k+t}, \tag{6}$$

where γ is the discount factor that controls the relative importance of future versus immediate rewards. We set $\gamma = 0.95$. The agent's objective is to learn an optimal policy π^* that maximizes the expected return:

$$\pi^* = \arg\max_{\pi} \mathbb{E}_{\pi}[G_k]. \tag{7}$$

We used a deep Q-network (Mnih et al., 2015), an instance of value-based reinforcement learning methods, to find the optimal behavior of metacognitive control π^*.

Model Setup and Simulations

To generate model predictions for each of the three experiments, we set up the list and reward structure to be identical to the experiment that we aim to model. Parameters at the object level describe constraints of cognitive resources, which we inherited from a previous implementation of resource-based memory models ($\tau = 0.072$, $r = 0.08$, $\theta_{\text{epi}} = 0.367$, $\sigma_{\text{epi}} = 0.256$; Ma et al., 2024). We then obtained how the model optimally allocates cognitive resources to each item (i.e., δ_k) when faced with different list and reward structures in a given experimental design. This model embodies our theoretical framework, predicting memory behavior as an optimal adaptation to varying reward environments, operating under a consistent set of cognitive constraints across all three experiments we aim to model (Figure 1A).

Experiment I

Many studies support that high-reward items are better remembered than low-reward items (Bowen et al., 2020; Gong & Li, 2014; Grandoit et al., 2024; Manga et al., 2020; Middlebrooks et al., 2017; Talmi et al., 2021). However, this memory advantage disappears when high- and low-reward items are studied in separate lists (Hellerstedt & Talmi, 2022; Hellerstedt et al., 2023; Talmi et al., 2021). In Experiment I we sought to explain this effect by proposing that participants adaptively allocate limited cognitive resources during memory encoding in a metacognitive process. We compared our model simulations with behavior patterns reported by Talmi et al. (2021). Our proposed model was set up based on the same rewards and trial structures as this experiment and was trained to maximize its overall reward gain. Because mixed and pure lists were randomized, participants were not informed in advance of each upcoming list type and only knew whether the list presented up to a given moment was mixed or pure. This information on the list type was also presented to the model. We then compared the recall patterns predicted by the model with those of human participants.

Method

Participants

Twenty-nine participants (aged 18–21) were recruited for an in-person study (Experiment 1 in the study by Talmi et al., 2021). Participation was compensated with course credit. The study was approved by the University of Manchester Research Ethics Committee.

Stimuli

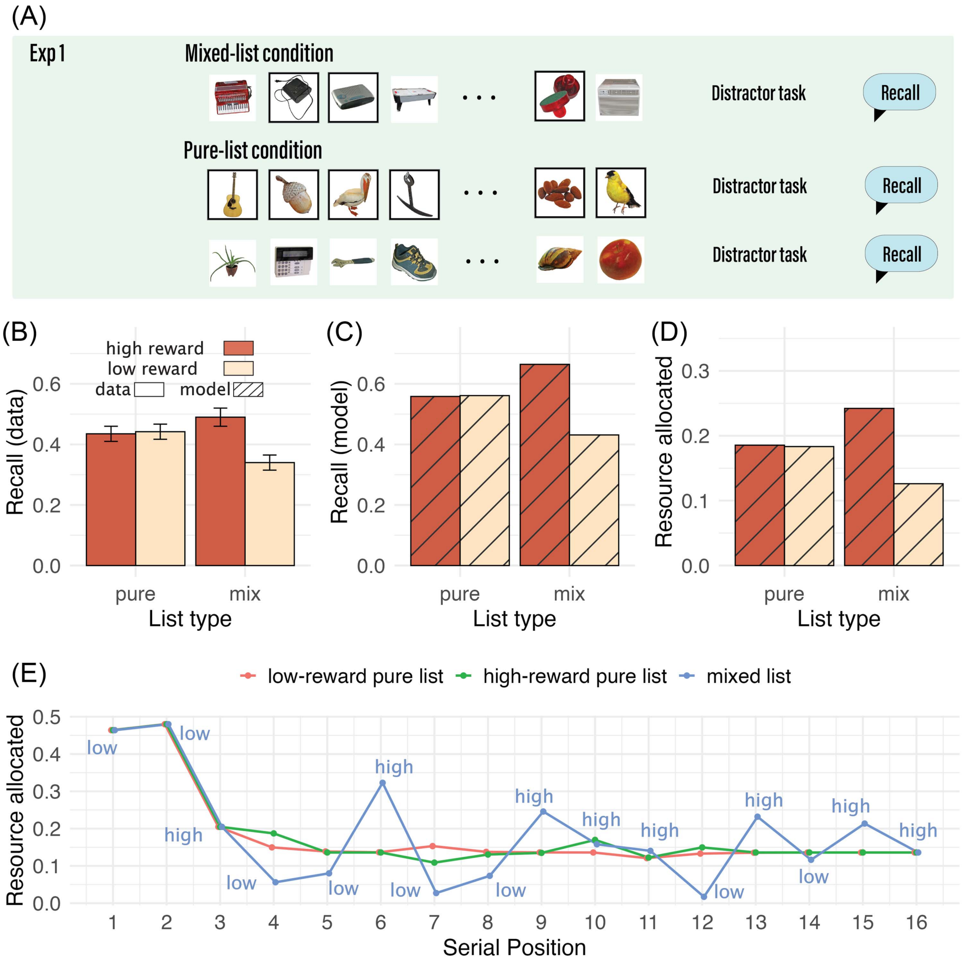
Stimuli consisted of 96 pictures drawn from two semantic categories: clothing and stationery. Pictures from the two categories were matched for familiarity and visual complexity. In each trial, participants viewed 16 pictures, with eight from each category. Images were either framed with a gray square (high reward) or left unframed (low reward). As depicted in Figure 2A, each participant viewed six trials, including two pure high-reward lists (all framed, e.g., HHHHHH), two pure low-reward lists (all unframed, e.g., LLLLLLL), and two mixed lists (half framed, half unframed, e.g., LHLHHL). Image assignments and the order of trials were randomized across participants.

Procedure

Participants completed a free recall task (Figure 2A). In each trial, they were shown 16 pictures sequentially. Each picture was presented for 2 s, followed by a white screen for 3.5, 4, or 4.5 s (randomized). Participants were not informed in advance about the reward condition of each upcoming list (pure or mixed list) or the

Figure 2

Experiment I: Enhanced Memory for High-Reward Items in Mixed Lists but Not in Pure Lists



Note. (A) Participants studied and recalled lists of pictures of low (unframed) or high rewards (framed). High-reward items are better remembered than low-reward items only in the mixed-list condition, but not in the pure-list condition, as shown for both (B) behavioral data reproduced from Figure 2A in Talmi et al. (2021) and (C) model predictions. Consistent with recall probability, more resources are allocated to high-reward than low-reward items in the mixed-list condition, but not in the pure-list condition, shown for both (D) model predictions and (E) an illustrative trial simulated from the model. Note that the model behavior was derived from optimal adaptation to the reward environment and not directly fitted to the data it seeks to explain. See the online article for the color version of this figure.

number of framed pictures they would see. Immediately after picture presentation, they completed a 60-s distractor task consisting of arithmetic questions on a sheet of paper in front of them. Following the distractor task, participants had 3 min to describe the pictures they remembered in any order. Participants were told that, for each list, they would receive £1 for each correctly recalled framed picture and 10 pence for each correctly recalled unframed picture. To be counted as correct, each recalled picture had to include at least three accurate descriptive details. More details of the experimental procedure can be found in Talmi et al.'s (2021) study.

Transparency and Openness

While our model reproduced the qualitative behavioral patterns reported by Talmi et al. (2021), we do not have access to the Experiment I data set. The model analysis code for this experiment can be accessed at https://osf.io/y6cqt/.

Results

We observed an alignment between model predictions (Figure 2C) and human recall patterns (Figure 2B) in pure lists versus mixed lists. Specifically, high-reward items were better recalled than low-reward items in mixed lists, but not in pure lists. Our model captures these effects through adaptive allocation of cognitive resources (Figure 2D): In mixed lists, high- and low-reward items compete directly for limited resources, with high-reward items prioritized over low-reward items (see an illustrative trial simulated from the model in Figure 2E). By contrast, in pure lists, competition occurs between items with the same rewards, so there is no clear advantage in allocating more resources to some items over others. As the total amount of available resources remains the same across both high-reward and low-reward pure lists, items within each list receive comparable resources during encoding (see Figure 2E). The alignment between model predictions and human recall patterns supports our hypothesis that participants adaptively allocate limited cognitive resources during memory encoding based on the overall reward environment. There is a tendency to allocate more cognitive resources to items carrying higher rewards; however, the exact amount of resources allocated to each item also depends on the overall reward environment. It is worth noting that our goal was to test the model's ability to predict qualitative behavior across different experiments, without directly fitting its parameters to each specific experiment. The model's overall higher performance compared with human participants (see Figure 2B and 2C) reflects specific parameters inherited from a previous study (Ma et al., 2024).

Experiment II

In Experiment I we showed that memory for an item depends not only on its own reward value but also on the overall reward environment it is situated (whether in a pure-list or mixed-list condition). Our model can explain these effects as it captures how items with different rewards compete for limited resources during encoding in a metacognitive process. The goal of Experiment II was to further investigate these mechanisms by examining how memory for a specific item is affected by other list items across more fine-grained reward magnitudes, analyzing a shared data set from Middlebrooks et al.'s (2017) study. Experiment II also serves as a test of the ability of our model to generalize to a new reward environment. Importantly, we maintained all model aspects and parameters from Experiment I changing only the reward environment to which the model must adapt optimally. In this experiment, participants were recruited to memorize lists of 20 words (Figure 3A). Each word was given a reward value ranging from 1 to 10 points. Participants were asked to recall the words in any order and maximize their total points. We then compared the recall patterns predicted by the model with those of human participants. If items compete for limited resources during encoding, as proposed by the metacognitive process of our model, we would expect an item's reward value to influence not only its own recall performance but also that of temporally nearby items in a study list.

Method

Participants

A total of 288 participants (aged 18–30) were included from two experiments originally reported by Middlebrooks et al. (2017). All participants were recruited at the University of California, Los Angeles, and received course credit for their participation. In each experiment, participants were randomly assigned to either a full-attention condition or one of several divided-attention conditions that imposed concurrent cognitive demands during encoding. For the current analysis, we include only participants in the full-attention condition across both experiments. This subset consisted of 70 participants.

Stimuli

Each participant in the full-attention condition completed six trials, with each trial containing 20 unique words randomly selected from a word bank of 280 nouns and verbs. Words ranged from four to seven letters in length. Each word was randomly paired with a point value from 1 to 10, with two words assigned to each value per list.

Procedure

Participants were shown six 20-word trials, with each word assigned a value between 1 and 10 points (Figure 3A). Words were presented sequentially for 3 s each. Immediately after each list presentation, participants were asked to recall as many words as possible in any order, with the goal of maximizing their total point value. A feedback screen displayed their score after each recall phase. More details of the experimental procedure can be found in Middlebrooks et al.'s (2017) study.

Transparency and Openness

Experiment II data were analyzed from Middlebrooks et al.'s (2017) study. The data and analysis code for this experiment can be accessed at https://osf.io/y6cqt/.

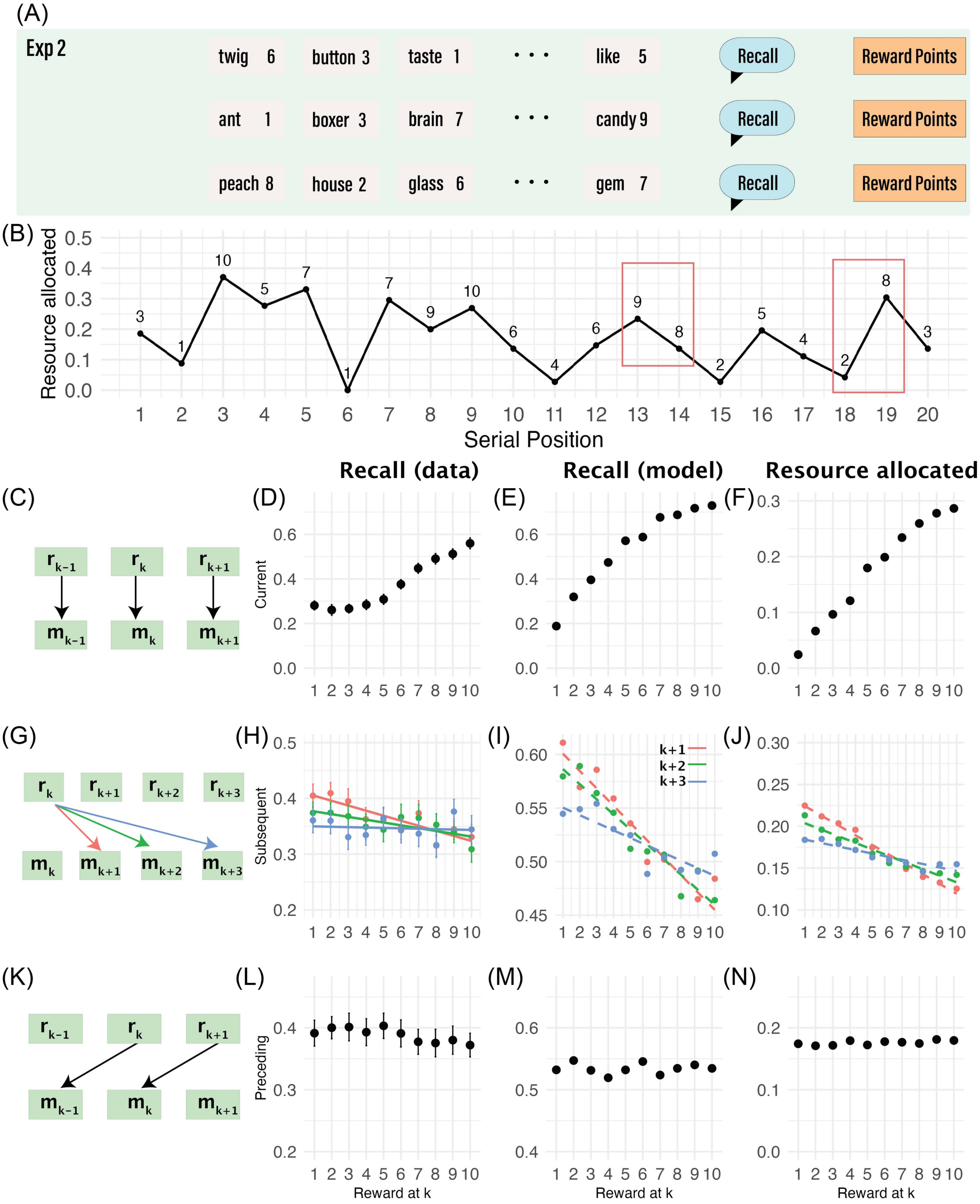
Results

Our model predictions align well with human data, revealing key patterns in the adaptive control of cognitive resources during encoding. We first examined how an item's reward value affects its own recall performance. Specifically, we analyzed the influence of reward at position k (r_k) on the recall probability of the word at the same position (m_k; Figure 3C). As shown in Figure 3D, participants had an increasingly better recall performance as the reward value increased—linear mixed-effects model: $\beta = .036$, $SE = 0.002$, $t(629) = 17.27$, $p < .01$—consistent with our model predictions (Figure 3E). The increasingly better recall performance can be explained by increasingly more resources allocated to item encoding (Figure 3F). This pattern can also be visualized from an illustrative trial simulated from the model (Figure 3B), where items with more reward points are generally allocated with more resources. This result reproduces the same effect in the mixed-list condition in Experiment I (Figure 2B), where high-reward items were better remembered than low-reward items. Our model successfully captures these effects regardless of the granularity of reward levels.

Although multiple accounts in the past can explain why high-reward items are better remembered than low-reward items, our model makes unique predictions on how reward affects the memory performance of temporally nearby items in a study list. Specifically, our model predicts that while a higher reward r_k for an item increases the recall probability for that item m_k (Figure 3E), it decreases the recall probability for the subsequent item m_{k+1} (Figure 3I). This prediction results from the competition among items for limited resources during encoding: Once resources are depleted, they require time to recover. When more cognitive resources are depleted in encoding an item with a higher reward at position k (Figure 3F), fewer resources remain available to encode the subsequent item at

Figure 3

Experiment II: Competition for Limited Cognitive Resources Among Temporally Nearby Items of Different Rewards



Note. (A) Participants studied and recalled lists of words of different reward points in experiments conducted by Middlebrooks et al. (2017). (B) An illustrative trial was simulated from the model, showing the amount of cognitive resources allocated to each item as a function of serial position and reward points associated with each item. Reward magnitude (from 1 to 10) has a positive effect on the recall probability and the amount of resources allocated for the current item (C–F), a negative effect on the subsequent items (G–J), and no effect on the preceding item (K–N), shown for both behavioral data (second column) and model predictions (last two columns). Note that the model behavior was derived from optimal adaptation to the reward environment and not directly fitted to the data it seeks to explain. See the online article for the color version of this figure.

position $k + 1$ (Figure 3J), resulting in lower recall probability at position $k + 1$ as reward increases (Figure 3I). This pattern can also be visualized from the illustrative trial simulated from the model (Figure 3B), where the item at Serial Position 14 (with 8 points) following a high-reward item with 9 points is allocated with fewer resources than the item at Serial Position 19 (with 8 points) following a low-reward item with 2 points. Our model further predicts that this sequential effect of reward will be less pronounced for subsequent items at position $k + 2$ and position $k + 3$ (see Figure 3I–3J), as a more distant subsequent item has more time to recover the cognitive resources.

Furthermore, although reward at position k affects the recall probability of the subsequent items (forward direction; Figure 3I), our model predicts that it leaves the recall probability of the preceding item at position $k - 1$ unchanged (backward direction; Figure 3M). No reward effect was shown in the backward direction because resources that have already been allocated to the item at position $k - 1$ cannot be reallocated or released, even if a high-reward item is subsequently encountered at position k. Consistent with this interpretation, our model shows that reward at position k does not modulate the amount of resources allocated to items at position $k - 1$ (Figure 3N).

Behavioral patterns align well with our model predictions, with an effect of r_k on m_{k+1} and m_{k+2} in the forward direction (Figure 3H), $k + 1$: $\beta = -.009$, $SE = 0.002$, $t(629) = -4.931$, $p < .01$; $k + 2$: $\beta = -.005$, $SE = 0.002$, $t(629) = -2.613$, $p < .01$; $k + 3$: $\beta = -.001$, $SE = 0.002$, $t(629) = -0.351$, $p = .726$, but no effect of r_k on m_{k-1} in the backward direction (Figure 3L), $\beta = -.003$, $SE = 0.002$, $t(629) = -1.653$, $p = .099$. We also confirmed that the forward effect of reward (position $k + 1$) was significantly stronger than the backward effect (position $k - 1$), as there was a significant interaction between direction and reward in a linear mixed-effects model, $\beta = -.007$, $SE = 0.004$, $t(681) = -2.08$, $p = .038$. While multiple accounts in the past can explain why high-reward items are better remembered than low-reward items, the effect of reward in the forward direction and a lack of effect in the backward direction, along with evidence from Experiment I, are unique evidence supporting our proposed model. These results support our hypothesis that memory behavior is shaped by the interaction between limited cognitive resources available during encoding and the specific reward environment.

Experiment III

In this experiment, we introduced a novel manipulation and collected a new data set to provide a strong test for our proposed metacognitive model. In Experiment II we showed that rewards affect temporally nearby items but only in the forward direction. The effect of reward on nearby items is due to the competition of limited cognitive resources, and the forward direction of this effect is a result of the nature of the cognitive resources; once allocated to an item, it cannot be reallocated even if a high-reward item is encountered at the subsequent position. While the reward of an item was not known to participants in Experiment I and Experiment II until the item was presented, we designed a reward environment in Experiment III wherein participants were informed of the reward structure of an entire list ahead of its presentation. We hypothesize that with the knowledge of upcoming rewards when encoding an item at position $k - 1$, participants can adaptively adjust or reserve their resource allocation at position $k - 1$ based on the reward at position k, therefore giving rise to competition of items in the backward direction. This would provide a strong test for our proposed account of reward effects on memory. As participants had limited time to learn the reward values for an entire list, we simplified the reward structure for participants to learn and used short lists consisting of six words per list with at most one switch in reward values midlist: HHHHHH, HHHLLL, LLLHHH, and LLLLLL (see Figure 4A). We then examined how reward values of the first or second half of the list influence the recall performance of the current half, the preceding half, and the subsequent half of the list. It is analogous to those conducted in Experiment II but simplified to analyses over halves of a list instead of over individual items in a list.

Method

Participants

One hundred twenty-five participants (aged 18–40) were recruited via the Prolific platform. They were all fluent English speakers and consented to participate in the study. Following our preregistered exclusion criteria, we excluded participants who failed the attention check, quit halfway, or reported using external help or indicating a lack of diligence in completing the task. Seventy-one participants remained in our analysis.

Stimuli

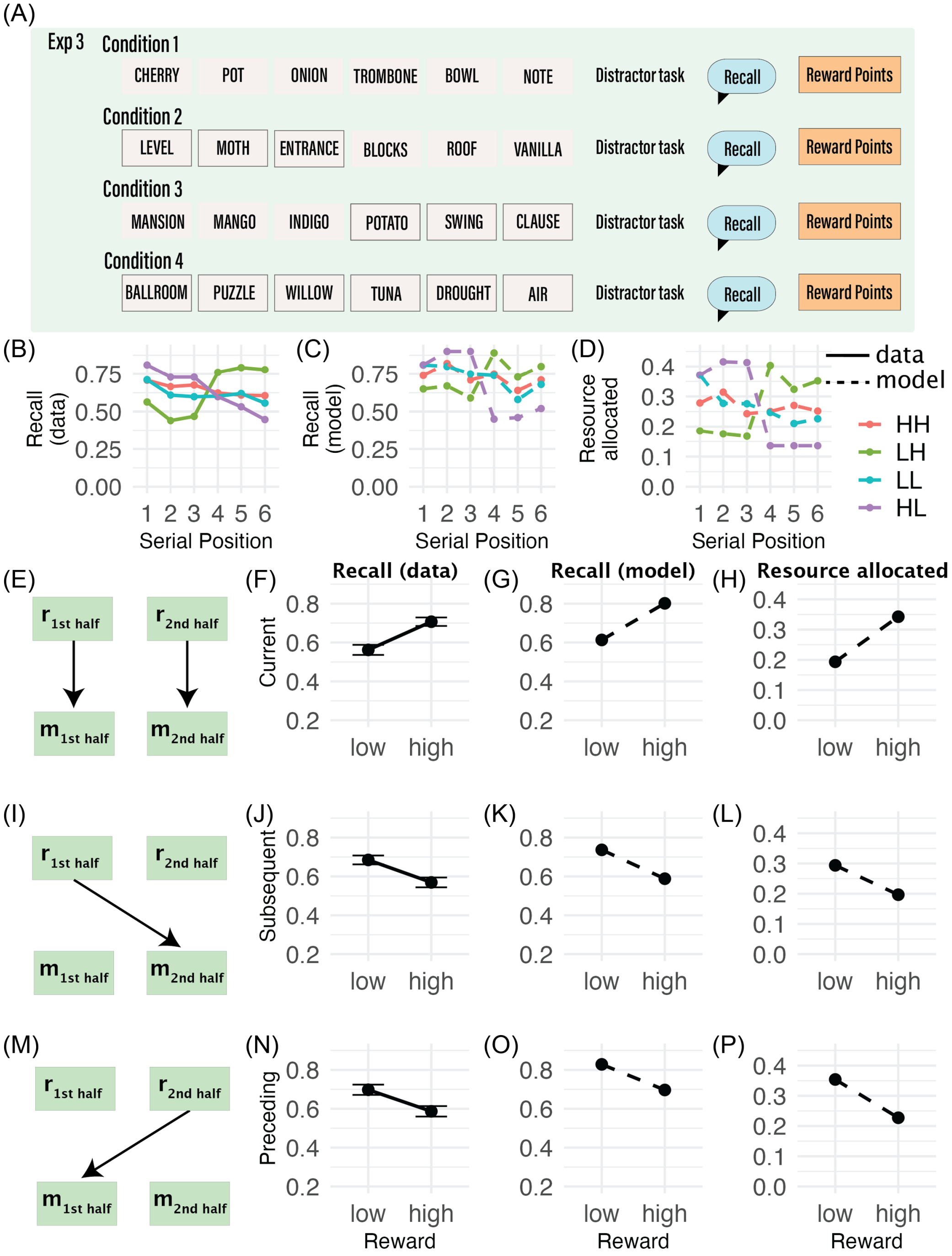
All word trials were randomly selected from a prior word pool of 326 words (Polyn et al., 2011). Each trial contained six words, each belonging to a different semantic category (Polyn et al., 2011). Some words were displayed with a gray rectangular frame around them, indicating a high-reward value (3 points). Other unframed words carried a low-reward value (1 point). The experiment included four types of reward structures: HHHHHH, HHHLLL, LLLHHH, and LLLLLL. Participants were informed of the reward structure before each list presentation. Each participant completed 17 trials in total, including one practice trial at the beginning and 16 experimental trials (four trials for each reward structure). The order of experimental trials was randomized for each participant.

Procedure

Participants completed a free recall task. In each trial, they were presented with six words, displayed sequentially. Each word appeared on the screen for 2 s, followed by a 0.5-s blank screen. Words were either framed or unframed, with framed words awarding high rewards (3 points) and unframed words awarding low rewards (1 point). At the beginning of each trial, participants were informed of the reward structure for the upcoming list. As each word was presented, they performed a size judgment task by pressing “Q” if the word was smaller than a shoebox or “P” if it was larger. To receive any points for a list, participants had to correctly classify more than half of the words in the list. This task was designed as an attention check. Following the list presentation, participants completed a 12-s distractor task, during which they solved three math problems in the form of “A + B + C = ?” Bonuses were awarded based on their performance. After the distractor task, participants had 15 s to recall the words from the just-presented list in any order. Their objective was to maximize their total reward

Figure 4

Experiment III: Reservation of Cognitive Resources When There Is Anticipation of High-Reward Items in the Near Future



Note. (A) Participants studied and recalled lists of words of low (unframed) or high rewards (framed). Importantly, participants were informed of the reward structure ahead of the list presentation. Rewards affect recall probability across serial positions for both (B) behavioral data and (C) model predictions, consistent with (D) the amount of resources allocated to items in the model. More specifically, rewards have a positive effect on the recall probability and the amount of resources allocated for the current items (E–H), a negative effect on the subsequent items (I–L), as well as the preceding items (M–P), shown for both behavioral data (second column) and model predictions (last two columns). Note that the model behavior was derived from optimal adaptation to the reward environment and not directly fitted to the data it seeks to explain. See the online article for the color version of this figure.

points, which influenced their final payment. The experiment was implemented using PsiTurk and Heroku and lasted approximately 25 min.

Transparency and Openness

We preregistered the study design, data exclusion criteria, model predictions, and all analyses in Experiment III and recruited participants via the Prolific platform (https://aspredicted.org/jyx6-mdqh.pdf). The data and analysis code can be accessed at https://osf.io/y6cqt/.

Results

Figure 4B and 4C compare behavioral data and model predictions of the recall probability across different serial positions for each list condition. The amount of resources allocated to items encoded at different serial positions is presented in Figure 4D. We can see that across data and model (Figure 4B–4D), recall probability of a given half of a list and its allocated resources depend not only on the reward values of the given half but also on the reward values of the other half of the list. To further unpack this observation, we analyzed the effect of reward values on current (Figure 4E), subsequent (Figure 4I), and preceding (Figure 4M) items following a similar approach to the analyses conducted in Experiment II. Recall probabilities were averaged for each half of the list. Consistent with the results of Experiment II, we found that while high-reward items were better remembered than low-reward items (Wilcoxon signed-rank test: $V = 252.5$, $n = 68$, $P < .01$, two-tailed; Figure 4F), subsequent items following high-reward items were associated with worse memory performance compared with those following low-reward items (Wilcoxon signed-rank test: $V = 1854$, $n = 66$, $P < .01$, two-tailed; Figure 4J). These results align with recall probabilities predicted from the model (Figure 4G and 4K) and the resources allocated (Figure 4H and 4L), supporting our hypothesis that participants adaptively allocate limited cognitive resources during memory encoding based on reward levels.

While rewards affect current and subsequent items (Figure 4F and 4J) in a manner comparable to Experiment II (Figure 3D and 3H), there are differences in how rewards affect preceding items between these two experiments. Experiment II did not exhibit an effect of reward on preceding memory performance; this is because resources cannot be reallocated to a high-reward item at position k once they have already been used to encode another item at position $k - 1$ (a key assumption in the object level). However, the above result could be reversed if one decides not to allocate resources at position $k - 1$ in the first place. A strong test for our proposed model is to examine if participants reserve resources at position $k - 1$ if they anticipate high-reward items in the upcoming future. In the current experiment, participants were informed of (so was the model) the reward structure of a list prior to the list presentation. We analyzed how rewards of the second halves influence memory performance of the first halves (Figure 4M). Consistent with our model predictions on the amount of resource allocation (Figure 4P) and recall performance (Figure 4O), items preceding high-reward items were associated with worse memory performance compared with those preceding low-reward items (Wilcoxon signed-rank test: $V = 1842.5$, $p < .01$, two-tailed; Figure 4N). This is in contrast to those in Experiment II wherein there was no effect of reward in the backward direction (Figure 3L), $\beta = -.003$, $SE = 0.002$, $t(629) = -1.653$, $p = .099$. Taken together, Experiment III provides a nontrivial and rigorous test of our hypothesis. The design difference between Experiment II and Experiment III was intentionally kept as a simple extreme of complete uncertainty versus complete predictability of the list reward structure. The resulting reversal of effects in the backward direction simultaneously validates two key components of our proposed model. First, it supports the role of cognitive constraints: that resources, once allocated, cannot be immediately released (as seen in Experiment II). Second, it confirms the role of adaptive behavior: that it is rational to reserve resources for an expected high reward (as seen in Experiment III).

General Discussion

Human episodic memory adaptively encodes important information. A fundamental question concerns how different reward values regulate this selective memory encoding process. In this work, we present a computational model that considers how people optimally allocate limited cognitive resources during memory encoding. Unlike previous accounts, which directly associate higher rewards with proportionally stronger memory encoding (Talmi et al., 2021; Zhou et al., 2023), we allow our model to adaptively decide how much to encode for each item based on the overall reward environment and one's limited cognitive resources. In three experiments, we showed that human behavior is consistent with our model predictions. In Experiment I our model accounts for the seemingly nonintuitive results that high-reward items are better remembered than low-reward items only when they are presented together, but not separately. Experiment II further investigated these mechanisms by examining how fine-grained reward magnitudes affect not only the current items but their temporally nearby items as predicted by the model. In Experiment III we collected a new data set to provide a strong test for our proposed model, demonstrating the role of reward anticipation in reversing an effect that we found in Experiment II. Model components and parameters were held the same across all three experiments, changing only the reward environment to which the model must adapt optimally. Together, our proposed theoretical framework provides a rational account that unifies diverse findings on the role of rewards on memory. We now turn to the broader implications of these results.

Our proposed framework posits that memory behavior modulated by rewards can be predicted by specifying an individual's cognitive constraints (i.e., limited cognitive resources during encoding) and their current reward environment. Assuming individuals are rational in optimizing their rewards, the framework generates unique predictions for memory behavior in any given reward environment. We have validated this model across three distinct reward environments, demonstrating its potential for broader generalizability. This approach follows the tradition of rational analysis, which views memory phenomena not as idiosyncratic outcomes but as optimal adaptation to the task environment (Anderson, 1990). Early work of rational analysis demonstrated that forgetting behavior commonly observed in laboratory settings is not a weakness or peculiarity of the memory system but reflects how the memory system sensibly adapts to the statistics of the environment (Anderson & Milson, 1989). Similarly, various findings that were originally interpreted as indicative of bias in human memory have been reexamined and instead found to reflect rational behavior in achieving people's goals

(Hemmer & Steyvers, 2009; Huttenlocher et al., 2000; Wilson et al., 2021; Xu et al., 2024). Building upon this foundation, recent research has extended the concepts of rational analysis by explicitly incorporating the constraints imposed by the limited computational resources (Gershman et al., 2015; Griffiths et al., 2015; Howes et al., 2009; Lewis et al., 2014). This framework, sometimes known as resource-rational analysis, has proven influential in offering explanations for a diverse range of psychological findings (Lieder & Griffiths, 2020). For instance, although accurate recall is the explicit goal of a memory task, people do not spend infinite effort and time searching their memory; instead, they rationally balance memory performance against retrieval effort, terminating searches when success becomes unlikely and prioritizing more accessible memories (Callaway et al., 2024). Furthermore, recent work has also begun to incorporate constraints from detailed models of memory to examine the emergence of optimal memory behavior (Lu et al., 2022; Zhang et al., 2023).

One of the crucial steps of the resource-rational analysis is to evaluate optimal solutions against empirical data, and if there are discrepancies, one should consider refining the model's assumptions regarding the cognitive constraints and reiterating or stopping until the model's assumptions are sufficiently realistic (Anderson, 1990; Griffiths et al., 2015). Accurately specifying the cognitive constraints is important, as different constraints can lead to different optimal behaviors even for the same task. However, iteratively adjusting constraints solely to match observed data risks becoming an exercise in curve fitting rather than achieving a genuine explanation. This challenge is particularly acute in memory research, as we are still actively identifying the precise nature and limits of cognitive constraints in memory, such as working memory capacity, attentional limits, or the specific cognitive operations permitted during encoding and recall. Consequently, although it is a fundamental question to understand how human memory is adaptive and supports goal-directed behavior—especially in the context of rewards, our ability to rigorously compare human behavior with optimal behavior is often limited by our incomplete knowledge of these underlying cognitive constraints.

Our present work circumvents this limitation by leveraging recent theoretical advances that characterize cognitive constraints during memory encoding (Popov & Reder, 2020; Reder et al., 2000, 2007). Crucially, these developments allow us to define cognitive constraints based on independent measurements and findings from prior research. This enables us to generate unique predictions about optimal memory behavior without engaging in the iterative refinement of constraints and optimization cycles. When specifying the cognitive constraints in our proposed framework, we inherited not only the exact model formulation but also model parameters from a previous study (Ma et al., 2024). Specifically, each time an item is stored in long-term memory, its encoding depletes a proportion of the currently available resources. The success of encoding is proportional to the resources allocated, and the depleted resources require time to recover. This model assumption has been validated over numerous behavioral findings (Kowialiewski et al., 2021; Mzrak & Oberauer, 2021; Oberauer, 2022; Popov & Reder, 2020; Popov et al., 2019, 2022). For example, it explains why memory performance suffers more after encoding difficult (e.g., low-frequency) items compared with easy (e.g., high-frequency) items—because the ability to encode the following item is impaired with less resources remaining. This effect diminishes with longer interstimulus intervals because there is more time for the resources to recover. The resource-based assumption also accounts for serial position effects, where performance declines across a list as initially maximal resources are gradually depleted. Beyond behavioral findings, recent computational modeling work linked these cognitive resources to neural substrates using electroencephalogram data (Ma et al., 2024). This work demonstrated that neural patterns associated with successful subsequent recall correspond to greater available resources during encoding compared with patterns for forgotten items (Ma et al., 2024), providing a mechanistic account for the well-established subsequent memory effects in the neuroscience literature (Kim, 2011; Paller & Wagner, 2002). Together, these behavioral and neural evidence highlight the importance of cognitive resources in contributing to successful memory formation and provide a computational framework to simulate these processes. Importantly, these prior theoretical and empirical developments establishing resource-based encoding constraints have not interacted with the literature of rewards on memory. Therefore, the alignment we observed between our model's predictions (i.e., optimized behavior under these independently validated constraints) and human memory behaviors provides a strong test for our proposed resource-rational account of reward-modulated memory and simultaneously provides independent validation for the specific resource-based constraints operating during encoding. Overall, our work bridges the literature identifying cognitive resources with the framework of resource-rational analysis, offering a principled explanation for how rewards shape memory behavior.

Our proposed model makes contact with several existing models of memory in explaining the effects of rewards or incentives. While our goal was to examine the adaptive allocation of cognitive resources during episodic memory encoding, a closely related model applying resource-rational analysis focuses on phenomena during visual working memory tasks (Van den Berg & Ma, 2018). Their model explains the set-size effect—where encoding precision declines as item count increases—as an optimal trade-off between behavioral performance and costly use of neural resources for encoding information. Specifically, as the set size grows, the decreasing likelihood of any single item being probed leads to fewer resources allocated to it. This framework assumes flexible total resource allocation, determined by a cost-benefit analysis; higher incentives can recruit more resources, albeit at a cost. This contrasts sharply with the inherently limited cognitive resources in our proposed resource-rational model. One possibility to reconcile these perspectives is that while an upper limit of resources may exist, it may be much higher than the average amount of resources needed to encode information into working memory in the tasks examined (Van den Berg & Ma, 2018). Therefore, a resource-rational model without assuming an upper limit in resources can still adequately capture behavior in those contexts. However, the processes of encoding information into episodic memory, like a long list of items presented sequentially, may be more likely to place the brain in a resource-depleted state, where the finite nature of the resource pool becomes a critical constraint. And under such a resource-depleted state, increasing rewards no longer increase memory performance (Talmi et al., 2021), where high-reward lists are equally remembered as low-reward lists. In fact, there is also evidence later in the working memory literature that shows no effect of rewards on the total amount of resources (Brissenden et al., 2023; Van den Berg et al., 2023), which poses challenges for the resource-rational model assuming flexible total resource allocation (Van den Berg & Ma, 2018). It remains a fruitful venue in the future to have an integrative framework to understand the role of adaptive allocation of resources during both working memory maintenance and long-term memory encoding.

Furthermore, previous studies have proposed alternative accounts to explain the finding that high-reward items are better recalled than low-reward items only in mixed lists, but not pure lists (Experiment I). For example, one account attributes this effect to item competition during retrieval (Geraci et al., 2013; McDaniel et al., 2005; Talmi et al., 2021), instead of competition of limited cognitive resources during encoding as proposed in our model. The emotional context maintenance and retrieval model posits that high-reward items always capture more attention during encoding, making them bind more strongly to their encoding context (Talmi et al., 2021). Recall performance of an item depends on how its encoding strength compares with that of other items at the time of retrieval. In mixed lists, high-reward items gain a retrieval advantage over low-reward items due to their stronger binding. In pure lists, however, all items share a similar binding strength, which removes the competitive advantage and causes the reward effect to disappear. Although a competition during retrieval could explain Experiment I (Talmi et al., 2021), it could not account for the finding that current rewards influence the recall of nearby items, which involves competition among items at encoding (Experiments II and III). Consistent with our proposed model account, an electroencephalogram study by Hellerstedt et al. (2023) demonstrated neural evidence of competition among items during encoding in mixed lists. We then consider accounts related to strategic encoding, such as the attentional borrowing account (Slamecka & Katsaiti, 1987). According to this account, learners would devote extra attention and selectively rehearse high-reward items, effectively borrowing encoding resources from other items in mixed lists. In this scenario, high-reward items receive an encoding boost and are recalled better, while low-reward items suffer. In pure lists, no single item can attract more attention, so the advantage disappears. Although this attention borrowing account is conceptually similar to our proposed model in explaining Experiment I, the lack of a computational implementation precludes specific predictions for Experiments II and III.

Our present work has focused on comparing model predictions and human data at the population level. Nonetheless, individuals differ in their reward sensitivity—that is, how participants vary in their selectivity for high versus low reward. Prior work has shown that older adults, despite recalling fewer items overall, remain as selective as younger adults when encoding high-reward items (Knowlton & Castel, 2022; Murphy et al., 2025). One possible mechanism that could contribute to such reward sensitivity emerges from the need to adapt one's resource allocation strategy to their specific level of cognitive constraint, such as how different individuals vary in their rate of resource recovery. Although further empirical studies are necessary, we demonstrated the plausibility of this mechanism in a model simulation (see Appendix A): When one's total cognitive resources are low, as induced by a slower recovery rate in older adults than young adults, it is rational to allocate more of the resources to high-reward items and compromise recall for low-reward items. Another possible mechanism that could contribute to the reward sensitivity is one's aversion to changing levels of resource allocation across items. Prior work has shown that human decision making reflects a trade-off between maximizing rewards and minimizing the complexity of the policy that maps states to actions (Gershman & Lai, 2021; Lai & Gershman, 2024). This policy complexity was found to suppress midbrain dopamine responses to reward outcomes, suggesting a neural cost to maintaining fine-grained strategies (Gershman & Lak, 2025). Relatedly, Piray and Daw (2021) found that when the brain needs to deviate from a default policy for more flexible, goal-oriented behavior, it incurs a cost. Building on these insights, we explored the possibility that participants with a higher aversion to policy complexity may prefer a simpler and less costly strategy wherein they allocate an equal proportion of cognitive resources to each item—giving rise to a lower reward sensitivity. To test this possibility, we extended our model by associating more complex encoding strategies with additional consumption of cognitive resources (see Appendix B). Our model simulations successfully predicted differences in reward sensitivity: Participants with a higher aversion to constantly changing resource allocation were associated with lower reward sensitivity. Taken together, our modeling framework has the potential to account for individual reward sensitivity, although further empirical work is needed to validate these model predictions. Our model also makes predictions on the role of the interstimulus interval. We demonstrate in model simulations that lengthening the interstimulus interval reduces the memory performance difference between high- and low-reward items (see Appendix C). This is because longer intervals allow cognitive resources to replenish before encoding the next item, thereby lessening the competition for cognitive resources among items with different rewards.

While we propose that rewards modulate memory encoding through adaptive resource allocation, it is important to acknowledge other related mechanisms through which rewards regulate memory encoding. For example, memory encoding is sensitive not only to the availability of cognitive resources but also to sudden shifts within a reward environment, which generates prediction errors. A surprising event characterized by a large reward prediction error—such as receiving an unexpectedly low reward after a series of high rewards—signals a significant change in the reward environment and tends to be better remembered (Rouhani et al., 2018). Enhanced memory for events with high prediction errors has been shown to result from the surprising event creating an event boundary during encoding, thereby reducing interference from memories formed prior to it (Rouhani et al., 2020). While our proposed model accounts for memory performance through a top-down metacognitive process, automatic processes also play critical roles in memory formation (Adcock et al., 2006; Knowlton & Castel, 2022) and can interact with the top-down processes (Murphy et al., 2022, 2024, 2025). For instance, an item's font size automatically influences our judgment of its importance while also affecting how participants selectively encode high-reward items (Murphy et al., 2024). Furthermore, people apply a habitual, serial processing during encoding (Murphy et al., 2022). Murphy et al. (2022) found that when words in a list are associated with different reward values, the influence of reward can reduce serial processing but does not entirely eliminate this more automatic process. It would be a fruitful future direction to integrate our proposed model with additional mechanisms of bottom-up processes. For instance, by integrating with a temporally drifting context during encoding (Polyn et al., 2009), our model could potentially be extended to account for this serial processing effect. Likewise, the proposed model could be integrated with other already-existing mechanisms that account for bottom-up effects like the font size effect. Beyond integrating bottom-up processes, our framework could potentially integrate with other factors such as those formalized in the priority-recall model (Castel et al., 2012). Specifically, the priority-recall model defined recall probability as the product of an item's priority value (the reward magnitude) and its time of recall value (the likelihood of an item being tested later). Our current model specifies the reward environment based on the reward magnitude for each item in the study list. Future work could incorporate a reward-discounted variable, which would assume that an item's reward is discounted during encoding by the likelihood that it will be tested later. Another factor that could influence how cognitive resources are allocated during memory encoding is item difficulty. Extending the model to environments where both reward and difficulty vary would allow for richer predictions about how people strategically allocate limited cognitive resources under competing task demands. Finally, our current model assumes cognitive resources are fully recovered before each new list. This simplification means the model cannot capture potential carryover effects across lists, where a demanding, high-reward list might impair performance on subsequent lists. Future studies could examine whether different lists compete for limited cognitive resources in the same way items do within a single list. While current results focus on studies of free recall, we expect the way cognitive constraints and rewards shape memory to generalize broadly. Because prior work has shown that cognitive resources are crucial for recognition and cued-recall tasks as well (Popov & Reder, 2020), we expect that the same interplay between rewards and cognitive constraints will shape performance in those paradigms too.

In summary, we present a model of metacognitive control for memory encoding to understand the role of reward on memory. It consists of a meta-level component that monitors the memory state and adaptively controls encoding strength for each piece of information. The model is constrained by limited cognitive resources available at encoding and optimally adapts to a given task environment to maximize total rewards. It successfully explains the previously observed nonintuitive finding in Experiment I as well as revealing novel effects of rewards on memory in Experiments II and III. Taken together, our results provide compelling evidence that reward-modulated memory encoding is an adaptive process sensitive to the overall reward environment rather than a passive response to individual item rewards. These findings offer a theoretical foundation for optimally assigning incentives to study materials in educational settings, which should consider both the value and the overall structure of the rewards.

References

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507–517. https://doi.org/10.1016/j.neuron.2006.03.036

Allen, R. J., & Ueno, T. (2018). Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Attention, Perception, & Psychophysics*, *80*, 1731–1743. https://doi.org/10.3758/s13414-018-1543-6

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press. https://doi.org/10.4324/9780203771730

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*(4), 703–719. https://doi.org/10.1037/0033-295X.96.4.703

Bowen, H. J., Gallant, S. N., & Moon, D. H. (2020). Influence of reward motivation on directed forgetting in younger and older adults. *Frontiers in Psychology*, *11*, Article 1764. https://doi.org/10.3389/fpsyg.2020.01764

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brissenden, J. A., Adkins, T. J., Hsu, Y. T., & Lee, T. G. (2023). Reward influences the allocation but not the availability of resources in visual working memory. *Journal of Experimental Psychology: General*, *152*(7), 1825–1839. https://doi.org/10.1037/xge0001370

Callaway, F., Griffiths, T. L., Norman, K. A., & Zhang, Q. (2024). Optimal metacognitive control of memory recall. *Psychological Review*, *131*(3), 781–811. https://doi.org/10.1037/rev0000441

Castel, A. D. (2007). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering. *Psychology of Learning and Motivation*, *48*, 225–270. https://doi.org/10.1016/S0079-7421(07)48006-9

Castel, A. D., McGillivray, S., & Friedman, M. C. (2012). Metamemory and memory efficiency in older adults: Learning about the benefits of priority processing and value-directed remembering. In *Memory and aging* (pp. 245–264). Psychology Press.

Castel, A. D., Murayama, K., Friedman, M. C., McGillivray, S., & Link, I. (2013). Selecting valuable information to remember: Age-related differences and similarities in self-regulated learning. *Psychology and Aging*, *28*(1), 232–242. https://doi.org/10.1037/a0030678

Cohen, M. S., Rissman, J., Suthana, N. A., Castel, A. D., & Knowlton, B. J. (2014). Value-based modulation of memory encoding involves strategic engagement of fronto-temporal semantic processing regions. *Cognitive, Affective, & Behavioral Neuroscience*, *14*, 578–592. https://doi.org/10.3758/s13415-014-0275-x

Geraci, L., McDaniel, M. A., Miller, T. M., & Hughes, M. L. (2013). The bizarreness effect: Evidence for the critical influence of retrieval processes. *Memory & Cognition*, *41*(8), 1228–1237. https://doi.org/10.3758/s13421-013-0335-4

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. https://doi.org/10.1126/science.aac6076

Gershman, S. J., & Lai, L. (2021). The reward-complexity trade-off in schizophrenia. *Computational Psychiatry*, *5*(1), 38–53. https://doi.org/10.5334/cpsy.71

Gershman, S. J., & Lak, A. (2025). Policy complexity suppresses dopamine responses. *Journal of Neuroscience*, *45*(9), Article e1756242024. https://doi.org/10.1523/JNEUROSCI.1756-24.2024

Gong, M., & Li, S. (2014). Learned reward association improves visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 841–856. https://doi.org/10.1037/a0035131

Grandoit, E., Cohen, M. S., & Reber, P. J. (2024). Reward enhancement of item-location associative memory spreads to similar items within a category. *Cognition and Emotion*, *38*(8), 1180–1195. https://doi.org/10.1080/02699931.2024.2352184

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229. https://doi.org/10.1111/tops.12142

Hellerstedt, R., Bekinschtein, T., & Talmi, D. (2023). Can neural correlates of encoding explain the context dependence of reward-enhanced memory? *Psychophysiology*, *60*(9), Article e14322. https://doi.org/10.1111/psyp.14322

Hellerstedt, R., & Talmi, D. (2022). Reward does not modulate forgetting in free recall tests. *Learning & Memory*, *29*(12), 430–434. https://doi.org/10.1101/lm.053631.122

Hemmer, P., & Steyvers, M. (2009). A Bayesian account of reconstructive memory. *Topics in Cognitive Science*, *1*(1), 189–202. https://doi.org/10.1111/j.1756-8765.2008.01010.x

Hennessee, J. P., Patterson, T. K., Castel, A. D., & Knowlton, B. J. (2019). Forget me not: Encoding processes in value-directed remembering. *Journal of Memory and Language*, *106*, 29–39. https://doi.org/10.1016/j.jml.2019.02.001

Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, *116*(4), 717–751. https://doi.org/10.1037/a0017187

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, *129*(2), 220–241. https://doi.org/10.1037/0096-3445.129.2.220

Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fmri studies. *Neuroimage*, *54*(3), 2446–2461. https://doi.org/10.1016/j.neuroimage.2010.09.045

Knowlton, B. J., & Castel, A. D. (2022). Memory and reward-based learning: A value-directed remembering perspective. *Annual Review of Psychology*, *73*(1), 25–52. https://doi.org/10.1146/annurev-psych-032921-050951

Kowialiewski, B., Lemaire, B., & Portrat, S. (2021). How does semantic knowledge impact working memory maintenance? Computational and behavioral investigations. *Journal of Memory and Language*, *117*, Article 104208. https://doi.org/10.1016/j.jml.2020.104208

Lai, L., & Gershman, S. J. (2024). Human decision making balances reward maximization and policy compression. *PLOS Computational Biology*, *20*(4), Article e1012057. https://doi.org/10.1371/journal.pcbi.1012057

Lewis, R. L., Howes, A., & Singh, S. (2014). Computational rationality: Linking mechanism and behavior through bounded utility maximization. *Topics in Cognitive Science*, *6*(2), 279–311. https://doi.org/10.1111/tops.12086

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, Article e1. https://doi.org/10.1017/S0140525X1900061X

Lu, Q., Hasson, U., & Norman, K. A. (2022). A neural network model of when to retrieve and encode episodic memories. *eLife*, *11*, Article e74445. https://doi.org/10.7554/eLife.74445

Ma, S., Popov, V., & Zhang, Q. (2024). A neural index reflecting the amount of cognitive resources available during memory encoding: A model-based approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *51*(3), 350–370. https://doi.org/10.1037/xlm0001364

Manga, A., Vakli, P., & Vidnyanszky, Z. (2020). The influence of anticipated monetary incentives on visual working memory performance in healthy younger and older adults. *Scientific Reports*, *10*(1), Article 8817. https://doi.org/10.1038/s41598-020-65723-5

McDaniel, M. A., Dornburg, C. C., & Guynn, M. J. (2005). Disentangling encoding versus retrieval explanations of the bizarreness effect: Implications for distinctiveness. *Memory & Cognition*, *33*(2), 270–279. https://doi.org/10.3758/BF03195316

Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017). Selectively distracted: Divided attention and memory for important information. *Psychological Science*, *28*(8), 1103–1115. https://doi.org/10.1177/0956797617702502

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Murphy, D. H., Hoover, K. M., Castel, A. D., & Knowlton, B. J. (2025). Memory and automatic processing of valuable information in younger and older adults. *Aging, Neuropsychology, and Cognition*, *32*(1), 142–168. https://doi.org/10.1080/13825585.2024.2360226

Murphy, D. H., Rhodes, M. G., & Castel, A. D. (2024). The perceived importance of words in large font guides learning and selective memory. *Memory & Cognition*, *52*(7), 1463–1476. https://doi.org/10.3758/s13421-024-01555-2

Murphy, D. H., Schwartz, S. T., & Castel, A. D. (2022). Serial and strategic memory processes in goal-directed selective remembering. *Cognition*, *225*, Article 105178. https://doi.org/10.1016/j.cognition.2022.105178

Mzrak, E., & Oberauer, K. (2021). What is time good for in working memory? *Psychological Science*, *32*(8), 1325–1337. https://doi.org/10.1177/0956797621996659

Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 26, pp. 125–173). Academic Press. https://doi.org/10.1016/s0079-7421(08)60053-5

Noh, E., Herzmann, G., Curran, T., & de Sa, V. R. (2014). Using single-trial eeg to predict and analyze subsequent memory. *NeuroImage*, *84*, 712–723. https://doi.org/10.1016/j.neuroimage.2013.09.028

Oberauer, K. (2022). When does working memory get better with longer time? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(12), 1754–1774. https://doi.org/10.1037/xlm0001199

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*(2), 93–102. https://doi.org/10.1016/s1364-6613(00)01845-3

Piray, P., & Daw, N. D. (2021). Linear reinforcement learning in planning, grid fields, and cognitive control. *Nature Communications*, *12*(1), Article 4942. https://doi.org/10.1038/s41467-021-25123-3

Polyn, S. M., Erlikhman, G., & Kahana, M. J. (2011). Semantic cuing and the scale insensitivity of recency and contiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 766–775. https://doi.org/10.1037/a0022475

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129–256. https://doi.org/10.1037/a0014420

Popov, V., Marevic, I., Rummel, J., & Reder, L. M. (2019). Forgetting is a feature, not a bug: Intentionally forgetting some things helps us remember others by freeing up working memory resources. *Psychological Science*, *30*(9), 1303–1317. https://doi.org/10.1177/0956797619859531

Popov, V., & Reder, L. M. (2020). Frequency effects on memory: A resource-limited theory. *Psychological Review*, *127*(1), 1–46. https://doi.org/10.1037/rev0000161

Popov, V., So, M., & Reder, L. M. (2022). Memory resources recover gradually over time: The effects of word frequency, presentation rate, and list composition on binding errors and mnemonic precision in source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(9), 1263–1280. https://doi.org/10.1037/xlm0001072

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 294–320. https://doi.org/10.1037//0278-7393.26.2.294

Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J., & Dickison, D. (2007). Experience is a double-edged sword: A computational model of the encoding/retrieval trade-off with familiarity. *Psychology of Learning* *and Motivation*, *48*, 271–312. https://doi.org/10.1016/S0079-7421(07)48007-0

Roberts, W. A. (1972). Free recall of word lists varying in length and rate of presentation: A test of total-time hypotheses. *Journal of Experimental Psychology*, *92*(3), 365–372. https://doi.org/10.1037/h0032278

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), 1430–1443. https://doi.org/10.1037/xlm0000518

Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020). Reward prediction errors create event boundaries in memory. *Cognition*, *203*, Article 104269. https://doi.org/10.1016/j.cognition.2020.104269

Sandry, J., & Ricker, T. J. (2020). Prioritization within visual working memory reflects a flexible focus of attention. *Attention, Perception, & Psychophysics*, *82*, 2985–3004. https://doi.org/10.3758/s13414-020-02049-4

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*(1), 1–20. https://doi.org/10.1146/annurev.ps.41.020190.000245

Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language*, *26* (6), 589–607. https://doi.org/10.1016/0749-596X(87)90104-5

Standing, L. (1973). Learning 10000 pictures. *Quarterly Journal of Experimental Psychology*, *25*(2), 207–222. https://doi.org/10.1080/14640747308400340

Sundby, C. S., Woodman, G. F., & Fukuda, K. (2019). Electrophysiological and behavioral evidence for attentional up-regulation, but not down-regulation, when encoding pictures into long-term memory. *Memory & Cognition*, *47*, 351–364. https://doi.org/10.3758/s13421-018-0871-z

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. A Bradford Book. https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

Talmi, D., Kavaliauskaite, D., & Daw, N. D. (2021). In for a penny, in for a pound: Examining motivated memory through the lens of retrieved context models. *Learning & Memory*, *28*(12), 445–456. https://doi.org/10.1101/lm.053470.121

Van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. *eLife*, *7*, Article e34963. https://doi.org/10.7554/eLife.34963

Van den Berg, R., Zou, Q., Li, Y., & Ma, W. J. (2023). No effect of monetary reward in a visual working memory task. *PLOS ONE*, *18* (1), Article e0280257. https://doi.org/10.1371/journal.pone.0280257

Wilson, S. A., Arora, S., Zhang, Q., & Griffiths, T. (2021). A rational account of anchor effects in hindsight bias. *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43, No. 43).

Xu, Z., Hemmer, P., & Zhang, Q. (2024). Towards a generalized bayesian model of reconstructive memory. *Computational Brain & Behavior*, *8*, 1–13. https://doi.org/10.1007/s42113-024-00222-8

Zhang, Q., Griffiths, T. L., & Norman, K. A. (2023). Optimal policies for free recall. *Psychological Review*, *130*(4), 1104–1124. https://doi.org/10.1037/rev0000375

Zhou, Z., Kahana, M. J., & Schapiro, A. C. (2023). *A unifying account of replay as context-driven memory reactivation*. bioRxiv. https://doi.org/10.1101/2023.03.22.533833

Murdock, Jr., B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482–488. https://doi.org/10.1037/h0045106

(*Appendices follow*)

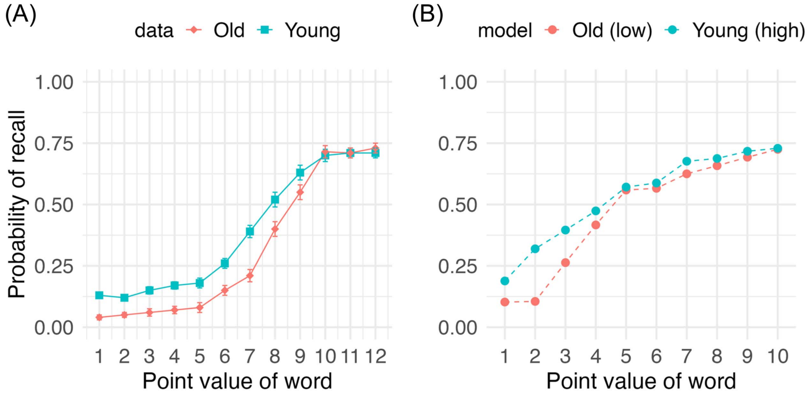
Appendix A

Resource Constraints as a Mechanism for Reward Sensitivity

While our model in the main text does not account for individual differences in reward sensitivity, we posit that several factors could contribute to this variability. One possible mechanism that could contribute to the reward sensitivity emerges from the need to adapt one's resource allocation strategy to their specific level of cognitive constraint, such as how fast cognitive resources recover. Prior work has shown that while older adults typically recall fewer words in free recall tasks, their ability to selectively encode high-reward items remains intact (Knowlton & Castel, 2022; Murphy et al., 2025). Specifically, the recall probability of high-reward items is similar between both age groups, but the recall probability of low-reward items is significantly lower for older adults (see Figure A1A; Knowlton & Castel, 2022). We hypothesize that variability in the resource recovery rate might account for this age difference in reward sensitivity. Specifically, older adults may have a slower cognitive resource recovery process, meaning that they require more time to recover the same amount of available resources compared with young adults. To test this hypothesis, we obtained the optimal encoding behavior of older adults using a smaller resource recovery rate while keeping all other parameters constant as the model presented in the main text. As shown in Figure A1B, the model simulation aligns with the behavioral data, indicating that an adaptation to different resource recovery rates could account for the observed differences between young and older adults. Specifically, a lower recovery rate leads to a reduced amount of total resources available during list encoding. To achieve a higher total reward, it is rational to maintain a higher resource allocation for the high-reward items and sacrifice the low-reward items rather than reducing the allocation equally for both low- and high-reward items.

Figure A1

Resource Constraints as a Mechanism for Reward Sensitivity



Note. (A) Recall probability as a function of the point value assigned to each word, reproduced from Figure 1B in Knowlton and Castel (2022). Older adults recalled high-reward words at levels comparable to young adults but showed reduced recall for low-reward words. (B) Recall probability as a function of the point value assigned to each word simulated from the model. Older adults are associated with lower recovery rate ($r = 0.07$), while young adults are associated with higher recovery rate ($r = 0.08$). Other parameters remain the same as the model presented in the main text ($\tau = 0.072$, $\theta_{epi} = 0.367$, $\sigma_{epi} = 0.256$). See the online article for the color version of this figure.

(*Appendices continue*)

Appendix B

Aversion to Policy Complexity as a Mechanism for Reward Sensitivity

Another factor that could contribute to reward sensitivity is the aversion to changing levels of resource allocation across items (we thank the reviewer for bringing this idea to us). Prior work has shown that human decision making reflects a trade-off between maximizing rewards and minimizing the complexity of the policy that maps states to actions (Lai & Gershman, 2024; Piray & Daw, 2021). A simple policy (e.g., uniform allocation of cognitive resources) avoids frequent switches in resource allocation, whereas a more complex policy adjusts allocation dynamically based on reward but incurs higher costs. Thus, individuals who are more averse to these costs may adopt simpler strategies and show lower reward sensitivity, while those who are less averse may exhibit greater sensitivity. To test our hypothesis, we consider that more complex strategies are associated with additional consumption of cognitive resources:

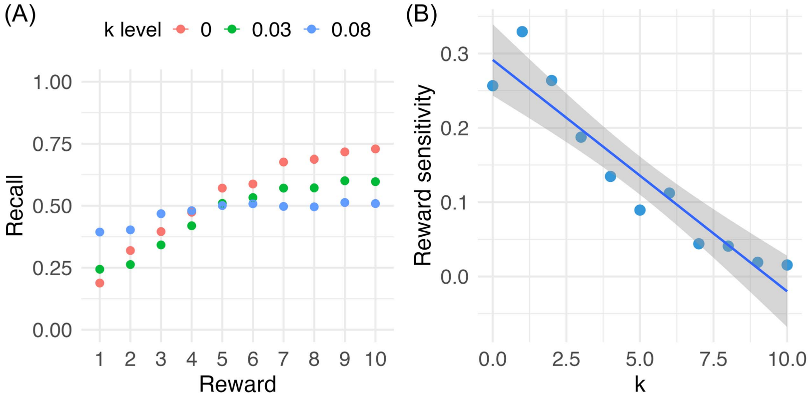
$$W_{\text{switch}, i} = |\delta_i - \delta_{i-1}| \times (W_i - W_{\text{sem}, i}) \times k \qquad \text{(B1)}$$

where δ_i and δ_{i-1} indicate the proportion of available resources allocated to encoding an item at position i and those allocated to the preceding item position $i - 1$ into episodic memory, and k is a scalar representing an individual's aversion to change in resource allocation for adjacent items ($|\delta_i - \delta_{i-1}|$). A larger k signifies a higher level of aversion, while a smaller k represents a lower level. Specifically, when $k = 0$, the model reverts to the original metacognitive model, where there is no cost incurred in constantly adapting the resource allocation level. When k is very large, the model is forced to allocate the same amount cognitive resources across all item positions. Other parameters represent the same concepts as in the main text. Specifically, W_i indicates the available resources at the beginning of encoding an item at position i, and $W_{\text{sem}, i}$ indicates the resources used for semantic processing of the item. The available resources before forming episodic memory are then updated to $W_i - W_{\text{sem}, i} - W_{\text{switch}, i}$. Keeping all other parameters the same as in the original model, we first simulated some examples to directly demonstrate how different values of k ($k = 0, 0.03, 0.08$) would influence memory performance. The reward environment is set to be the same as Experiment II. As shown in Figure B1A, a lower k, representing a smaller cost aversion, is associated with higher reward sensitivity.

We then simulated 11 different k values, ranging from 0 to 1 with a 0.1 interval, with each value representing a subject. We tested whether an individual's aversion to resource change (i.e., k) is associated with reward sensitivity. Reward sensitivity is measured as the slope of a linear mixed-effects model fitted to the recall performance as a function of the current reward. As shown in Figure B1B, there is a significant negative Pearson correlation between k and reward sensitivity, $r = 0.95$, $t(9) = -8.68$, $p < 0.01$. This finding supports our hypothesis that when an individual has a strong aversion to changes in their resource allocation, they exhibit a lower reward sensitivity.

Figure B1

Aversion to Policy Complexity as a Mechanism for Reward Sensitivity



Note. (A) Recall performance as a function of reward value under different levels of aversion, k. Larger values of k indicate stronger aversion, and as k increases, sensitivity to reward decreases. Other parameters remain the same as the model presented in the main text ($\tau = 0.072$, $\tau = 0.072$, $r = 0.08$, $\theta_{\text{epi}} = 0.367$, $\sigma_{\text{epi}} = 0.256$). (B) A negative correlation between k and reward sensitivity. See the online article for the color version of this figure.

(*Appendices continue*)

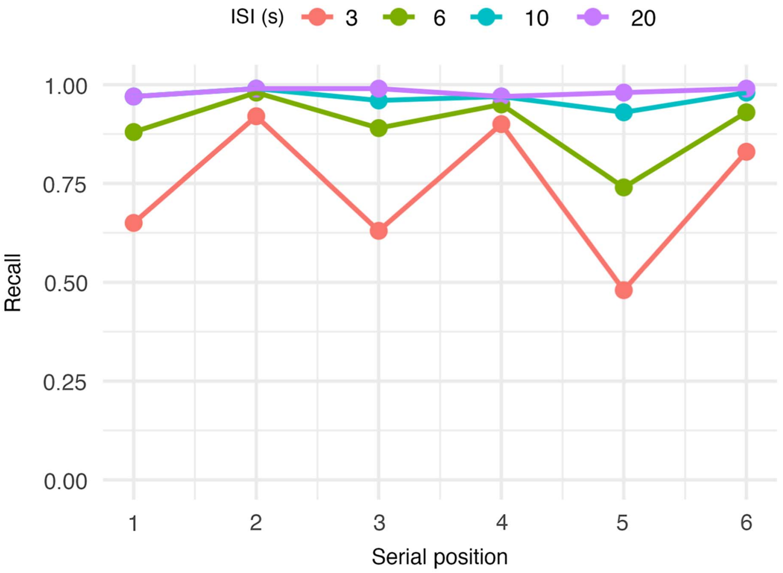
Appendix C

The Role of Interstimulus Interval on Reward-Enhanced Memory

Previous studies (Popov et al., 2022; Popov & Reder, 2020) have shown an interaction between memory performance and the interstimulus interval (ISI). While we are not aware of any empirical studies that have examined the effect of the ISI on the reward-enhanced memory, our model provides a clear prediction. We expect that as the ISI lengthens, cognitive resources would have more time to recover. This would result in a smaller difference in the amount of available resources at the beginning of encoding the next item, leading to a decreased memory advantage for high-reward items. If the ISI is long enough that allows for a full recovery of resources, our model predicts that there would be no difference in memory performance between high- and low-reward items. We used the reward structure LHLHLH and simulated the recall under different ISIs. As shown in Figure C1, the performance difference between high-reward and low-reward items continuously decreases with an increasing ISI, until it is eventually eliminated.

Figure C1

Recall Performance as a Function of Serial Position With Different Interstimulus Intervals



Note. Recall performance as a function of serial position with different interstimulus intervals (ISI) when the reward structure is LHLHLH. All parameters remain the same as the model presented in the main text ($\tau = 0.072$, $r = 0.08$, $\theta_{epi} = 0.367$, $\sigma_{epi} = 0.256$). See the online article for the color version of this figure.

Received July 17, 2025
Revision received October 22, 2025
Accepted November 20, 2025 ■