Selective memory for reward-relevant features is modulated by expertise during reward learning

Yirong Xiong (yirong.xiong@student.tuebingen.de)
University of Tübingen, Maria-von-Linden-Str. 6
Tübingen, Baden-Württemberg 72076 Germany

Nir Moneta (moneta@mpib-berlin.mpg.de)
Max Planck Institute for Human Development, Lentzeallee 94,
14195 Berlin, Germany

Mihály Bányai (mihaly.banyai@tuebingen.mpg.de)
Max Planck Institute for Biological Cybernetics
Max-Planck-Ring 8, 72076 Tübingen, Germany

Charley M. Wu (charley.wu@uni-tuebingen.de)
University of Tübingen, Maria-von-Linden-Str. 6
Tübingen, Baden-Württemberg 72076 Germany

Abstract

Efficiently prioritizing important information is crucial for human memory function. Previous studies have demonstrated that the value of stimuli can selectively influence memory, with humans selectively remembering reward-relevant information. Here, we add to this understanding by decomposing reward-relevance to different compositional features, which collectively define the value of a stimulus with differing importance. Using combined reward learning and recognition memory tasks operating on the same set of stimuli, we investigate the impact of feature importance on memory. Our findings suggest that selective memory for the most rewarding feature is influenced by the depth of expertise during reward learning. This research adds to a growing body of research on the mechanism of value-based memory, with novel insights into how expertise influences selective memory for reward-relevant features.

Keywords: Value-based Memory; Value-directed Remembering; Value Learning; Working Memory; Arbitary

Introduction

In an information-rich world, we are constantly bombarded with overwhelming amounts of data. However, the human memory system has limited capacity, making it impossible to remember everything in a lossless manner (Baddeley, 2012; Anderson & Hulbert, 2021). As a result, it is crucial to prioritize important information and allocate our limited cognitive resources efficiently (Cowan, 2000; Nagy, Török, & Orbán, 2020). For example, when you visit a rental apartment, you may remember the price and location, but not the color of the doors or tiles on the staircase. This is because we prioritize certain factors over others based on their significance.

Previous work has shown that the value associated with stimuli can selectively influence memory (Knowlton & Castel, 2022; Schultz, Stoffregen, & Benoit, 2023; Middlebrooks, Kerr, & Castel, 2017). Thus, we tend to remember rewarding information better than unimportant or less rewarding information (Thomas, FitzGibbon, & Raymond, 2015). This phenomenon can be explained by rate-distortion theory (Sims, 2016), which suggests that memory capacity can be adaptively allocated to minimize distortion (Gershman, 2021; Bates & Jacobs, 2020), where reward-relevant distortions may have a larger impact based on a cost-benefit trade-off (Bhui, Lai, & Gershman, 2021).

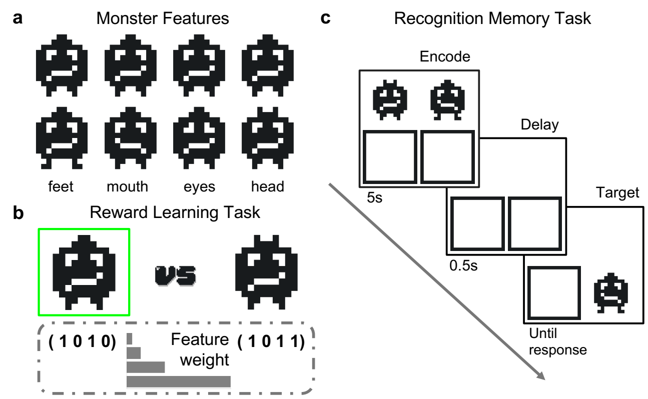

Figure 1: Value-based Working Memory Experiment. **a**) Monster stimuli. Each monster had four binary features, including feet, mouth, eyes and head. **b**) Top: Reward learning task. Participants needed to choose the "more powerful" monster in a series of fights between monsters. Each choice was followed by feedback (correct vs. incorrect). Bottom: The "power" (i.e., value) of each monster was determined by weighted sum of its features. Each feature had a different weight, with each feature type randomly assigned a value of either 1 or 0. **c**) Recognition memory task. Participants were shown two randomly sampled monsters on the upper row for 5 seconds. After a 0.5-second delay, a target monster was presented in one of the bottom row slows. Participants were asked to whether the target monster was identical to the monster directly above it.

Internal representations of values guide many daily decisions (De Martino & Cortese, 2023; Canas & Jones, 2010). While previous studies suggest that reward-relevant choices impact memory performance, it remains unclear whether this also applies to how people selectively remember individual features, which collectively define the value of a stimulus. For example, in the apartment rental scenario, the price and location may be more important than others. Therefore, the objective of this study is to investigate whether the ability to selectively remember stimuli is influenced by the hierarchy of feature importance from a reward learning task.

In this work, we use a combined reward learning and recognition memory task, both with the same set of stimuli. During reward learning, participants learned that features make differing contributions to reward. In the memory task, performed both before and after reward learning. Our results revealed systematic differences in feature-specific recall, where we found better selective memory for the most rewarding feature—but only in the best performing reward learners. This work provides important insights into the mechanisms of human memory and how we make the most out of limited cognitive resources.

Methods

We designed a Value-based Working Memory experiment, consisting of a reward learning task and a recognition memory task. Both tasks used the same monster stimuli (Fig. 1a), each defined by a set of four binary features with $2^4 = 16$ monsters in total. Each feature was comprised of six pixels and with two pixels marking the difference between each feature value. The reward learning task required participants to identify which pair of monsters was "more powerful" based on a weighted-sum of features (Fig. 1b), while the memory task had participants determine if a target monster was the same or different from the probe above (Fig. 1c).

Participants and Design. We recruited 100 participants on Prolific ($M_{age} = 34.26$; $SD = 12.68$; 49 female, 2 non-binary). The task took on average 36.27 $\pm$ 13.41 minutes and participants earned £8.58 $\pm$ £1.13 on average.

Materials and procedure. Participants were first given instructions about the stimuli and completed a comprehension check. They then performed 1 block of the memory task ("pre-learning"), 4 blocks of the reward learning task, and then 1 final block of the memory task ("post-learning").

In the **memory task**, participants completed 40 trials in each block. On each trial, they were shown 2 monsters for 5 seconds as a probe, followed by a target monster directly below. Participants were asked to respond 'Y' if the monster was the same as the one above, or 'N' if it was different. Participants received feedback after every 10 trials. Probe monsters were randomly sampled, while the target monster selectively modified one of the four features with $p(\text{change}|f_i) = .5$.

In the **reward learning task**, participants first completed an interactive tutorial, before starting the main task consisting of 240 trials across 4 blocks. On each trial, participants were shown a pair of monsters (sampled without replacement from all 240 pairwise combinations) and asked to predict which monster was "more powerful". Participants were

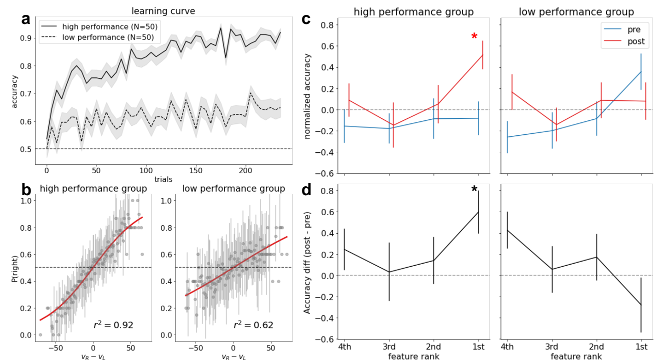

Figure 2: Results. **a**) Learning curve from the reward learning task. Participants are separated (high vs. low performance) based on a median-split of their accuracy on the last half of trials. **b**) Choice probability as a function of value difference between monsters (using last half of trials), separated by high vs. low performance. As the value-difference between monsters increases (x-axis), the likelihood of selecting the monster with the higher value also increases (y-axis). Participants in the high performance group (left) exhibit a sharper slope near the inflection point compared to the low performance (right). **c**) Memory performance in pre- and post-learning blocks. We first normalized accuracy w.r.t., to each participant's individual accuracy (y-axis), and the results across trials where each target feature was modified. Then we arranged the features based on their weight-rankings in the learning task (x-axis; highest weight on the right). The star indicate significant differences from change ($p = .001$). **d**) Memory difference in pre- and post-learning. The high performance group (left) showed a significant improvement in performance on the most important feature ($p = .004$), whereas the low performance group (right) did not demonstrate any significant changes in memory.

informed that the "power" of a monster is defined by their features, with some features being more important than others. We defined the power of each monster as a weighted sum of features $P(v) = \sum_i f_i w_i$, where the weights were the same for all participants $\mathbf{w} = [0.62, 0.24, 0.09, 0.03]$. Participants received truthful feedback after each trial and a summary of their performance every block. Feature values (e.g., which feet were assigned 0 or 1) and feature weights were randomly counterbalanced across participants.

Results

We first present analyses of the reward learning task, before looking at the pre- and post-learning effects on memory.

Learning performance

Overall, participants performed better than chance ($t_{99} = 16.37, p < .001$), with accuracy consistently improving over trials (Pearson's $r = .65, p < .001$; Fig. 2a). However, not all participants performed equally well, and we separated participants using a median split on their average accuracy in the last half learning trials, creating high and low performance groups ($N_{\text{high}} = N_{\text{low}} = 50$). We also find that accuracy can be described as a function of the value difference between the two monsters (Fig. 2b), where larger value differences increased the probability of choosing the higher-valued monster. This effect was stronger for the high performance group (Sigmoid fit: $r^2_{\text{high}} = .92$) than the low performance group

($r^2_{\text{low}} = .62$).

Selective memory

Next, we looked at memory performance, where overall, participants performed better than chance in both pre- (high: $t_{49} = 12.07, p < .001$; low: $t_{49} = 11.60, p < .001$) and post-learning memory blocks (high: $t_{49} = 14.34, p < .001$; low: $t_{49} = 11.19, p < .001$). Notably, the high performance group demonstrated better performance in the post-learning memory block compared to the non-learning group ($t_{98} = 2.52, p = .013$).

We then separated each trial based on the target feature that was modified, arranging them based on their feature weights in the reward learning task (Fig. 2c). We report the normalized accuracy, where 0 corresponds to the average accuracy of each participant across all memory trials where the target did not match the probe. Thus, positive values correspond to better than average memory performance, and negative values correspond to worse than average. These results reveal that the high performance group selectively remembered the highest weighted feature after exposure to the reward learning task ($t_{49} = 3.90, p = .001$), but not for other features (all $p > .05$). This selective memory effect was not found in the low performance group.

In Figure 2d, we report the difference in normalized accuracy from pre- to post-learning, where we find that memory for the highest weighted feature was significant improved, but only in the high performance group ($t_{49} = 3.01, p = .004$).

Conclusion

We investigated whether exposure to reward learning influenced selective memory of reward-predictive features. Our results reveal that the highest performing reward learners selectively remembered the most highly-weighted feature, but not for all participants. These findings highlight the adaptive selectivity of memory, which was modulated by the depth of expertise in reward learning. These results contribute to a growing body of research on value-based working memory and provide insights into how learning feature importance for reward prediction selectively influences memory.

Acknowledgements

This work is supported by the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC2064/1–390727645 and the Marie Skłodowska-Curie Action RELEARN-DLV-897042.

References

Anderson, M. C., & Hulbert, J. C. (2021). Active forgetting: Adaptation of memory by prefrontal control. *annual review of psychology*, *72*, 1–36.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual review of psychology*, *63*, 1–29.

Bates, C. J., & Jacobs, R. A. (2020). Efficient data compression in perception and perceptual memory. *Psychological Review*, *127*(5), 891–917. doi: 10.1037/rev0000197

Bhui, R., Lai, L., & Gershman, S. J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, *41*, 15–21.

Canas, F., & Jones, M. (2010). Attention and Reinforcement Learning: Constructing Representations from Indirect Feedback. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *32*(32).

Cowan, N. (2000). Processing limits of selective attention and working memory: Potential implications for interpreting. *Interpreting*, *5*, 117–146. doi: 10.1075/intp.5.2.05cow

De Martino, B., & Cortese, A. (2023, January). Goals, usefulness and abstraction in value-based choice. *Trends in Cognitive Sciences*, *27*(1), 65–80. doi: 10.1016/j.tics.2022.11.001

Gershman, S. J. (2021). The rational analysis of memory. , 19.

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. doi: 10.1146/annurev-psych-032921-050951

Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017, August). Selectively Distracted: Divided Attention and Memory for Important Information. *Psychological Science*, *28*(8), 1103–1115. doi: 10.1177/0956797617702502

Nagy, D. G., Török, B., & Orbán, G. (2020, October). Optimal forgetting: Semantic compression of episodic memories. *PLOS Computational Biology*, *16*(10), e1008367. doi: 10.1371/journal.pcbi.1008367

Schultz, H., Stoffregen, H., & Benoit, R. G. (2023, January). *A reward effect on memory retention, consolidation, and generalization?* PsyArXiv. doi: 10.31234/osf.io/89s2k

Sims, C. R. (2016, July). Rate–distortion theory and human perception. *Cognition*, *152*, 181–198. doi: 10.1016/j.cognition.2016.03.020

Thomas, P. M. J., FitzGibbon, L., & Raymond, J. E. (2015, November). Value conditioning modulates visual working memory processes. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(1), 6. doi: 10.1037/xhp0000144