A Normative Account of the Influence of Contextual Familiarity and Novelty on Episodic Memory Policy

Qihong Lu (qihong.lu@columbia.edu)
Center for Theoretical Neuroscience;
The Mortimer B. Zuckerman Mind Brain Behavior Institute;
Columbia University, New York, NY, USA

Kenneth A. Norman (knorman@princeton.edu)
Department of Psychology;
Princeton Neuroscience Institute;
Princeton University, Princeton, NJ, USA

Daphna Shohamy (ds2619@columbia.edu)
Department of Psychology;
The Mortimer B. Zuckerman Mind Brain Behavior Institute;
Kavli Institute for Brain Science;
Columbia University, New York, NY, USA

Abstract

How do humans decide when to retrieve and when to encode episodic memories (EMs)? Empirical results show that seeing a familiar stimulus biases subjects toward retrieval, while seeing a novel stimulus biases subjects toward encoding, even though that stimulus is incidental to the task. From a normative standpoint, it is unclear why the familiarity of incidental stimuli should bias EM. We hypothesized that these biases could arise because the EM policy – whether to retrieve or encode at a given moment – is learned in an environment where stimulus familiarity is autocorrelated in time. We present an EM-augmented neural network that learns an EM policy using reinforcement learning. Learning to encode was facilitated by allowing the reward obtained by retrieval to propagate back to reinforce the action of encoding this memory. As our model learns in an autocorrelated environment, empirically observed effects of familiarity emerged. This is because, in an environment where familiar stimuli tend to precede other familiar stimuli, familiarity indicates that relevant EMs are present, making retrieval more useful. Novelty encourages encoding for the same reason. Our results suggest that the influences of familiarity and novelty are adaptive features of human EM policy in response to autocorrelated environments.

Keywords: episodic memory (EM); familiarity & novelty; neural network; reinforcement learning (RL)

Prior work has shown that familiarity/novelty triggers temporally lingering biases towards episodic retrieval/encoding (Duncan, Sadanand, & Davachi, 2012; Duncan & Shohamy, 2016; Patil & Duncan, 2018; Duncan, Semmler, & Shohamy, 2019). These effects were attributed to familiarity reducing levels of acetylcholine, which alters the tendency of the hippocampus to enact pattern completion vs. pattern separation (Easton, Douchamps, Eacott, & Lever, 2012; Meeter, Murre, & Talamini, 2004; Hasselmo, Wyble, & Wallenstein, 1996). However, existing theories do not articulate why familiarity/novelty should trigger these biases from a normative standpoint.

We argue that these biases are adaptive features of human EM policy in response to the temporal autocorrelation in natural environments, where familiar/novel stimuli often precede other familiar/novel stimuli. This means that familiarity/novelty indicates the presence/absence of other relevant EMs, making episodic retrieval/encoding more useful. Through two simulations, we show that the empirically observed familiarity/novelty effects emerge as an agent learns an EM policy (of whether to retrieve or encode) in an autocorrelated environment. Our view resonates with the hypothesis that lingering mental states, in general, are adaptations to autocorrelated environments (Honey, Mahabal, & Bellana, 2023). Our work goes beyond prior models of how familiarity influences EM policy (Lu, Hasson, & Norman, 2022) by i) providing a mechanism to learn retrieval and encoding jointly and ii) examining how autocorrelation in the environment biases EM policy in different scenarios, thereby accounting for empirical data.

A neural network model of EM policy

Familiarity signal Inspired by prior works (Ji-An, Stefanini, Benna, & Fusi, 2023; Bogacz & Brown, 2003; Norman & O'Reilly, 2003), we use an autoencoder to compute familiarity/novelty (Fig.1a). For every experienced stimulus, the model takes a gradient descent step to minimize reconstruction mean squared error (MSE). Hence, experienced stimuli will have a lower MSE relative to novel stimuli, making MSE a noisy indicator of familiarity/novelty.

EM policy The EM policy network (Fig.1a) is a recurrent neural network that takes the familiarity signal as input and outputs a binary action – retrieve or encode, in line with prior work suggesting pattern completion and pattern separation cannot be executed simultaneously (O'Reilly & McClelland, 1994; Meeter et al., 2004; Hasselmo et al., 1996). The model learns a policy for retrieving/encoding that maximizes reward.

The tasks we considered involve choosing between two items (a or b) based on their values (see Fig.1b). When the model encodes, it stores the item and its value jointly. When the model retrieves, it looks for an exact item match in its EM buffer. If there is one, its value will be retrieved. Then, a neural network decision model (Fig.1a) takes the retrieved value and produces a binary output (a or b) indicating its choice. The reward is the value of the chosen item, which serves as the RL signal (Mnih et al., 2016) for both the EM policy and decision models. The model learned to recall when the item is familiar (Fig.1c) and choose the item with higher value (Fig.1e).

Credit assignment (CA) via retrieved memory Learning to encode is challenging as encoding never leads to immediate reward – an encoded memory does not affect the agent's behavior until it is retrieved later on. With regular temporal credit assignment (CA), if a memory was encoded T time steps ago, and retrieving it leads to a reward of r_t, the credit assigned to the encoding action is $r_t\gamma^T$, where $\gamma \in [0,1)$ is the discount factor. Because $r_t\gamma^T$ decreases exponentially as T increases, the learning signal for encoding actions is weak.

Our model overcomes this issue by performing CA via retrieved memory – when retrieval leads to r_t, the time point when this memory was encoded receives additional credit $r_t\gamma$; i.e., the act of encoding a memory is retrospectively reinforced when retrieving this memory leads to a reward (Fig.1b). This is cognitively plausible if we assume that, when the model retrieves the encoded item and value, it also retrieves the state it was (previously) in when it decided to encode; once this state is retrieved, the model can reinforce the link between this state and the encoding action.

Simulation 1. The context familiarity/novelty effect

In Duncan and Shohamy (2016) – hereafter, **DS16** – human subjects had to choose the higher-valued item from two new items or an old versus a new item (Fig.1b). All item values were sampled uniformly from the set $\{0, 0.25, 0.5, 0.75, 1\}$. As the value for each specific item was fixed, when an old item re-appears (e.g. the feather in Fig.1b), it is optimal to retrieve its value (0.75), then choose the old item if its value is higher than the mean (= 0.5), and choose the new one other-

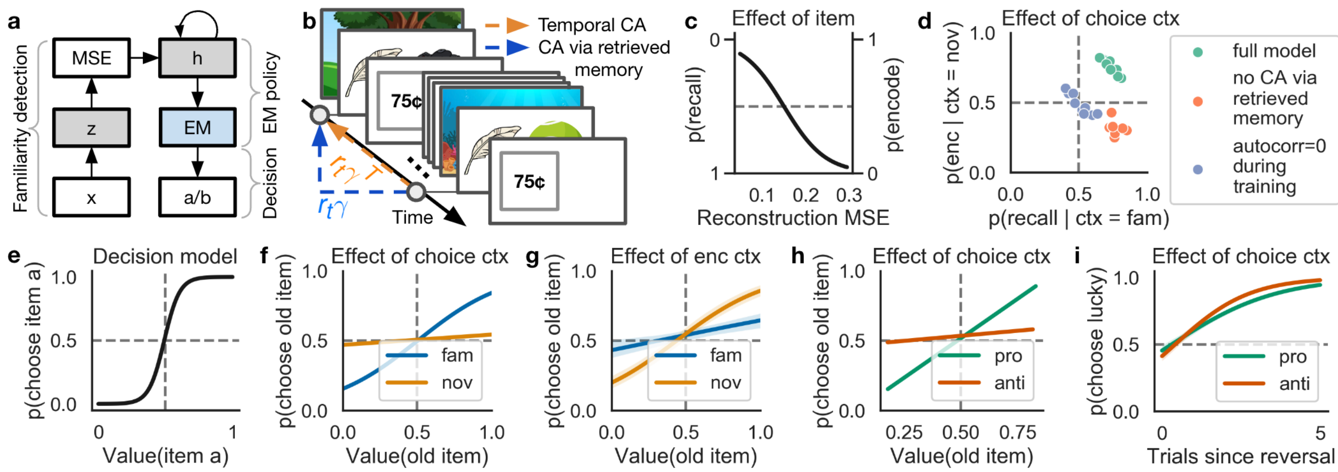

Figure 1: **a)** The model architecture. **b)** The DS16 task design, showing different credit assignment (CA) schemes. **c)** The model is more likely to recall/encode when item reconstruction MSE is low/high. **d)** The model is more likely to recall/encode when the context is familiar/novel. This is not true for models without CA via retrieved memory and models trained with zero autocorrelation in the environment. Every point is a model. **e)** The decision model chooses/avoids item a when its value is high/low. **f, g)** Model results for the DS16 task: In the model, the old item's value influenced choice more strongly (i.e., steeper slope) **f)** when the choice was made in a familiar context, and **g)** when the old item's value was encoded in a novel context. **h, i)** Model results for the D19 task: **h)** The old item's value influenced choice more strongly in the pro-episodic condition. **i)** After a reversal, the model switched to the lucky color more rapidly in the anti-episodic condition, indicating it is more dependent on incremental estimates. Error bands = 95% bootstrapped confidence interval. N=10 models. fam = familiar; nov = novel; ctx = context; enc = encode.

wise. Importantly, "context", an image presented immediately before choice, impacted subjects' use of EM – a familiar context made people's choice more dependent on retrieved value, and a novel context facilitated value encoding, even though context is incidental in this task and contains no information about item familiarity/novelty.

We hypothesized that these context familiarity/novelty biases are adaptive responses to environment auto-correlation. To test this, we trained our model on the DS16 task, in an environment where context familiarity and item familiarity were perfectly correlated. To simulate human data, we tested the model (weights frozen) when context and item familiarity were orthogonal, which is the case in the experiment. During choice, the model was indeed more likely to recall/encode when the preceding context was familiar/novel (Fig.1d). This allowed the model to capture the human data (see Fig.2 & 4 in DS16). Fig.1f,g shows the probability of choosing the old item as a function of its previously observed value – as in the human data, the slope, indicating the level of influence of retrieved value, was steeper i) when the preceding context was familiar (Fig.1f) and ii) when a novel context was presented at encoding, right before the value of the item was initially observed (Fig.1g). Models trained in an environment with zero auto-correlation did not show these biasing effects of context (Fig.1d). Models without CA via retrieved memory also did not show these effects (Fig.1d) – they were very unlikely to encode, which hindered learning.

Simulation 2. Incremental versus retrieved value

Duncan et al. (2019) – hereafter, **D19** – designed a setting where subjects had to coordinate incrementally tracked value versus retrieved value. The design is the same as DS16, except that items had two possible colors. At a given moment, one color is "lucky". The mean values for the lucky/unlucky color were 63/37. The lucky color reversed multiple times within an experiment session, so incrementally tracking the two mean values is useful. We augmented the decision model by computing exponentially weighted moving averages (EWMA) for the two colors. When an old item re-appears, it is optimal to retrieve its value, compare it with the EWMA estimate of the new item (based on its color), and choose the item with a higher value. To examine the effect of context familiarity/novelty, subjects were tested in 1) the pro-episodic condition, where context and item familiarity were perfectly correlated, versus 2) the anti-episodic condition, where context and item familiarity were perfectly anti-correlated. Concretely, in the anti-episodic condition, subjects always experienced i) a novel context before they had to recall (i.e., choose between old vs. new) and ii) a familiar context before they had to encode (i.e., choose between new vs. new). Empirically, the pro-episodic condition enhanced the influence of retrieved value on choice, whereas the anti-episodic condition enhanced people's dependency on incremental estimates.

We trained our model in the pro-episodic condition of the D19 task, where context familiarity and item familiarity are perfectly correlated, and then tested the model on both conditions. The model captured the human data (see Fig.4 in D19). First, EM influenced choices more strongly in the pro-episodic condition (Fig.1h). Additionally, in the anti-episodic condition, our model was faster at switching to the lucky color after a reversal, suggesting its choices were more dependent on incremental value estimates (Fig.1i).

To conclude, the empirically observed influences of familiarity/novelty can emerge as adaptive features of EM policy in response to auto-correlation in the natural environment.

Acknowledgments

We would like to thank Jonathan Nicholas, Kim Stachenfeld, and Stefano Fusi for their helpful feedback.

This work is supported by a NIMH/NIH grant (1R01MH130068) to D.S., a Multi-University Research Initiative grant to K.A.N.(ONR/DoD N00014-17-1-2961), a Zuckerman Institute Alan Kanzer Postdoctoral Fellowship to Q.L., Gatsby Foundation (GAT3708) and Kavli Institute for Brain Science at Columbia.

References

Bogacz, R., & Brown, M. W. (2003). Comparison of computational models of familiarity discrimination in the perirhinal cortex. *Hippocampus*, *13*(4), 494–524.

Duncan, K., Sadanand, A., & Davachi, L. (2012, July). Memory's penumbra: episodic memory decisions induce lingering mnemonic biases. *Science*, *337*(6093), 485–487.

Duncan, K., Semmler, A., & Shohamy, D. (2019, October). Modulating the use of multiple memory systems in value-based decisions with contextual novelty. *Journal of cognitive neuroscience*, *31*(10), 1455–1467.

Duncan, K., & Shohamy, D. (2016, November). Memory states influence value-based decisions. *Journal of experimental psychology. General*, *145*(11), 1420–1426.

Easton, A., Douchamps, V., Eacott, M., & Lever, C. (2012, November). A specific role for septohippocampal acetylcholine in memory? *Neuropsychologia*, *50*(13), 3156–3168.

Hasselmo, M. E., Wyble, B. P., & Wallenstein, G. V. (1996). Encoding and retrieval of episodic memories: role of cholinergic and GABAergic modulation in the hippocampus. *Hippocampus*, *6*(6), 693–708.

Honey, C. J., Mahabal, A., & Bellana, B. (2023, August). Psychological momentum. *Current directions in psychological science*, *32*(4), 284–292.

Ji-An, L., Stefanini, F., Benna, M. K., & Fusi, S. (2023, January). Face familiarity detection with complex synapses. *iScience*, *26*(1), 105856.

Lu, Q., Hasson, U., & Norman, K. A. (2022, February). A neural network model of when to retrieve and encode episodic memories. *eLife*, *11*, e74445.

Meeter, M., Murre, J. M. J., & Talamini, L. M. (2004). Mode shifting between storage and recall based on novelty detection in oscillating hippocampal circuits. *Hippocampus*, *14*(6), 722–741.

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T. P., Harley, T., ... Kavukcuoglu, K. (2016, February). Asynchronous methods for deep reinforcement learning.

Norman, K. A., & O'Reilly, R. C. (2003, October). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological review*, *110*(4), 611–646.

O'Reilly, R. C., & McClelland, J. L. (1994, December). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus*, *4*(6), 661–682.

Patil, A., & Duncan, K. (2018, January). Lingering cognitive states shape fundamental mnemonic abilities. *Psychological science*, *29*(1), 45–55.