Paradoxical replay maintains balanced and robust representations of task structure

Hung-Tu Chen (htchen.gr@dartmouth.edu)
Department of Psychological & Brain Sciences, Dartmouth College
Hanover, NH 03755, USA

Matthijs A. A. van der Meer (mvdm@dartmouth.edu)
Department of Psychological & Brain Sciences, Dartmouth College
Hanover, NH 03755, USA

Experience replay is a powerful mechanism to learn efficiently from limited experience. Despite several decades of striking experimental results, the factors that determine which experiences are selected for replay remain unclear. A particular challenge for current theories is that on tasks that feature unbalanced experience, rats paradoxically replay the less-experienced trajectory. To understand why, we simulated a feedforward neural network with two regimes: rich learning (structured representations tailored to task demands) and lazy learning (unstructured, task-agnostic representations). We find that rich, but not lazy, representations degrade following unbalanced experience, an effect that could be reversed with paradoxical replay. To test if this computational principle can account for the experimental data, we examined the relationship between paradoxical replay and learned task representations in the hippocampus. Strikingly, we find a strong association between the richness of learned task representations and the paradoxicality of replay. Taken together, these results suggest that paradoxical replay specifically serves to protect rich representations from the destructive effects of unbalanced experience.

Keywords: neural network models; rich and lazy learning; replay; hippocampus; splitter signal

Introduction. Leading theories of the content and function of hippocampal replay emphasize a dual role in both memory consolidation (typically during off-line states; a fundamentally retrospective process) and in planning upcoming decisions (on-line; a prospective process). Both these accounts predict that the content of hippocampal replay should align closely with proximal experience, whether recent or upcoming. However, several studies have found preferential replay of trajectories *not* recently experienced (Carey et al., 2019; Gillespie et al., 2021; Gupta et al., 2010). Interestingly, studies reporting such "paradoxical replay" have in common that experience is highly unbalanced (i.e. repeating one specific trajectory). A largely separate computational literature has shown that unbalanced training makes memory storage in neural networks susceptible to destructive interference (McClelland et al., 1995; Norman et al., 2005). This idea suggests a specific hypothesis for the function of paradoxical replay: to protect memories from the destructive effects of unbalanced experience. Here, we test this hypothesis in a model neural network and in neural data from the rat hippocampus.

Results. We first explore the effects of unbalanced experience in three-layer neural networks trained on a contextual discrimination task (context 1: L+R-, context 2: L-R+; Figure 1a, b) using two distinct training regimes with different computational tradeoffs (Flesch et al., 2022, 2023). Using a balanced training regime where all task conditions occur with equal frequency, "rich learning" networks (small initial weights variance, $\sigma = 0.0025$) learned more slowly compared to "lazy learning" networks (large variance, $\sigma = 0.25$). Input-to-hidden weights in rich networks captured the task structure into a low-dimensional representation (Figure 1c) robust to perturbation in the inputs. Conversely, in lazy networks, input weights remained unchanged and high-dimensional (Figure 1d) making the network more sensitive to input perturbations (data not shown).

Next, we examined the effects of unbalanced training on these two different network regimes. Under these conditions, rich networks developed biased representations, with input weights favoring positive outcomes (Figure 1e, top row) resulting in decreased performance on the task. We hypothesized that paradoxical replay could help rebalance representations for both outcomes. To test this, we simulated replays with an overrepresentation of the less-experienced task conditions after biased training (Figure 1f). This led to unbiased representations for both outcome contingencies (Figure 1f, top row). In contrast, lazy networks were insensitive to unbalanced training, and did not require rebalancing using paradoxical replay (Figure 1e, f, bottom row). Thus, rich vs. lazy learning networks were differentially sensitive to unbalanced training regimes, and paradoxical replay can rebalance the input weights of a rich learning network in the face of unbalanced training.

The computational principle unveiled above suggests a working hypothesis for the function of paradoxical replay in the rodent hippocampus: it serves to protect rich, but not lazy, task representations from interference due to unbalanced training. This hypothesis predicts that on tasks, subjects and sessions with low-dimensional "contextual" task representations characteristic of rich learning, paradoxical replay should be strong. In contrast, when high-dimensional, lazy representations are found, paradoxical replay should be weak.

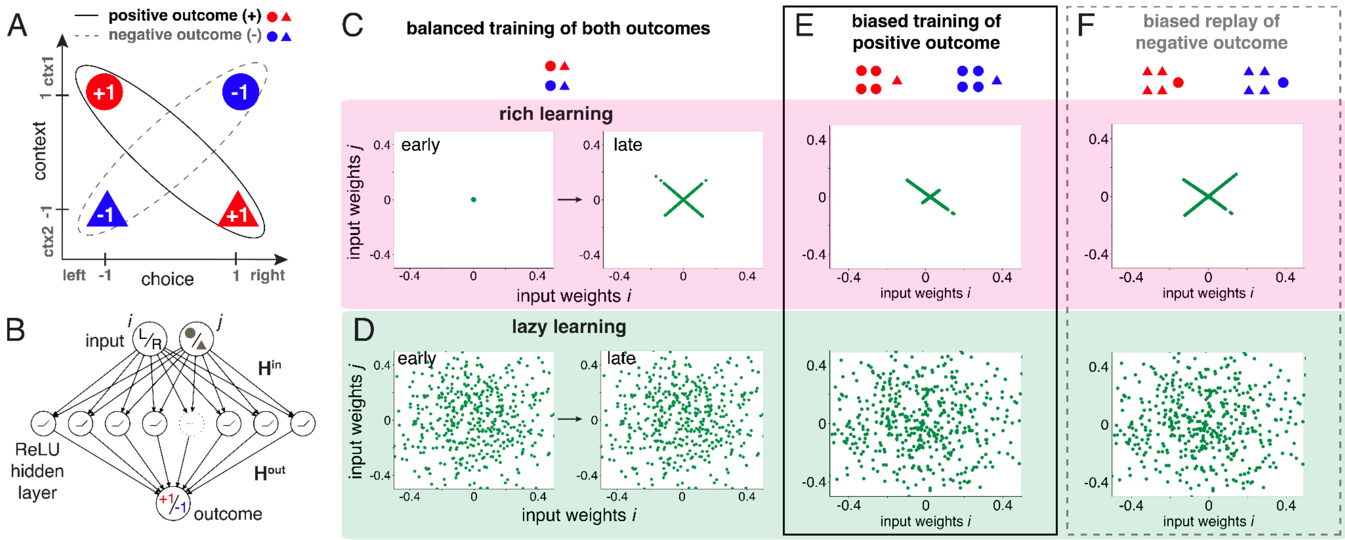

Figure 1: (A) Contextual discrimination task schematic: in context 1, left but not right is rewarded; these contingencies are reversed for context 2. **(B)** Feedforward neural network architecture trained on this task. **(C)** Two distinct learning regimes were initialized: rich-learning (*small* initial weight variances) and lazy-learning (*large* initial weight variances)**.** In rich-learning networks, input weights undergo notable changes, eventually learning the task structure. **(D)** Conversely, under the lazy learning regime, input weights stay unchanged during training. **(E)** Biased training on the positive contingency resulted in biased representations of input weights in rich-learning networks, while lazy-learning networks remained unaffected. **(F)** Paradoxical replay rebalanced the input weights of rich-learning networks, while lazy-learning networks remained unaffected.

To test this idea in experimental data, we first operationalize rich vs. lazy learning in terms of representational similarity of hidden unit activity: strong separation between hidden activity encoding the two contexts is indicative of low-dimensional rich learning, and conversely more mixed encoding indicates high-dimensional, lazy learning (Figure 2a). We then correlated this measure of rich learning, applied to hippocampal place cells, with the strength of paradoxical replay in the Carey et al. (2019) data set. Strikingly, those sessions in which rats showed the largest separation between the two trajectories, indicative of a rich representation, showed the strongest paradoxical replay (Figure 2b).

Discussion. These results support a normative explanation for paradoxical replay, suggesting it prevents rich representations from the detrimental impacts of unbalanced experience. We validate the core prediction that “richer” representations in the rodent hippocampus exhibit increased paradoxical replay compared to "lazy" representations. In doing so, our results link together for the first time two different experimental phenomena: “splitter cells” in the hippocampus (Duvelle et al. 2023) and replay. Moreover, our theory refines the notion of consolidation in complementary learning systems theory, in that representations from rich, but not lazy, task representations benefit from interleaving.

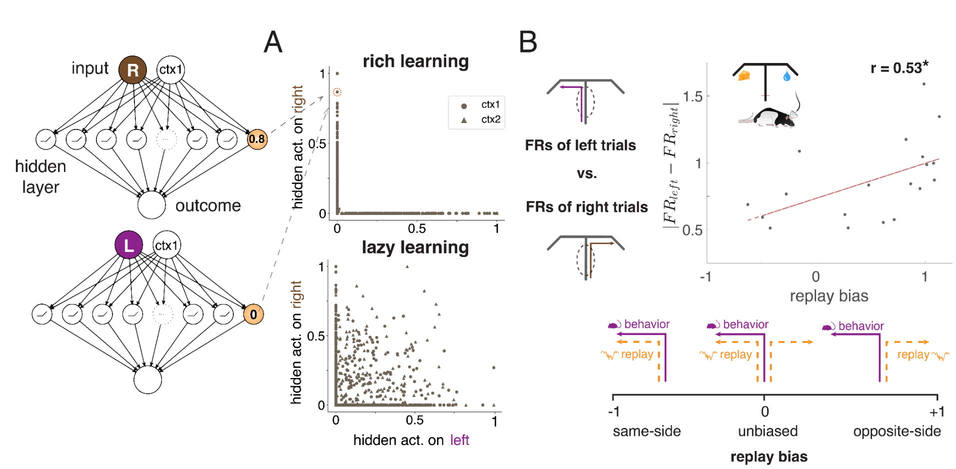

Figure 2: (A) Rich-learning networks exhibit selective activation ("splitting") for either left or right choices, while lazy-learning networks have mixed encoding of both choices. **(B)** In the Carey data those sessions with larger representational distance showed stronger paradoxical replay ($r = 0.53$, $p < 0.05$).

Acknowledgments

We thank Jeremy Manning, John Murray, James Fitzgerald and Cristina Savin for their helpful feedback and discussion. This work was supported by NIMH R01 MH123466 to MvdM.

References

Carey, A. A., Tanaka, Y., & van der Meer, M. A. A. (2019). Reward revaluation biases hippocampal replay content away from the preferred outcome. *Nature Neuroscience*, *22*(9), 1450–1459.

Flesch, T., Juechems, K., Dumbalska, T., Saxe, A., & Summerfield, C. (2022). Orthogonal representations for robust context-dependent task performance in brains and neural networks. *Neuron*, *110*(7), 1258–1270.e11.

Flesch, T., Saxe, A., & Summerfield, C. (2023). Continual task learning in natural and artificial agents. *Trends in Neurosciences*, *46*(3), 199–210.

Gillespie, A. K., Astudillo Maya, D. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., Roumis, D. K., Eden, U. T., & Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron*, *109*(19), 3149–3163.e6.

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., & David Redish, A. (2010). Hippocampal Replay Is Not a Simple Function of Experience. In *Neuron* (Vol. 65, Issue 5, pp. 695–705). https://doi.org/10.1016/j.neuron.2010.01.034

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457.

Norman, K. A., Newman, E. L., & Perotte, A. J. (2005). Methods for reducing interference in the Complementary Learning Systems model: oscillating inhibition and autonomous memory rehearsal. *Neural Networks: The Official Journal of the International Neural Network Society*, *18*(9), 1212–1228.