Rethinking dopamine as generalized prediction error: supplementary figures

Matthew P.H. Gardner[1], Geoffrey Schoenbaum[1,2,3], and Samuel J. Gershman[4]
[1]Intramural Research Program of the National Institute on Drug Abuse, NIH
[2]Department of Anatomy and Neurobiology, University of Maryland School of Medicine
[3]Department of Neuroscience, Johns Hopkins School of Medicine
[4]Department of Psychology and Center for Brain Science, Harvard University

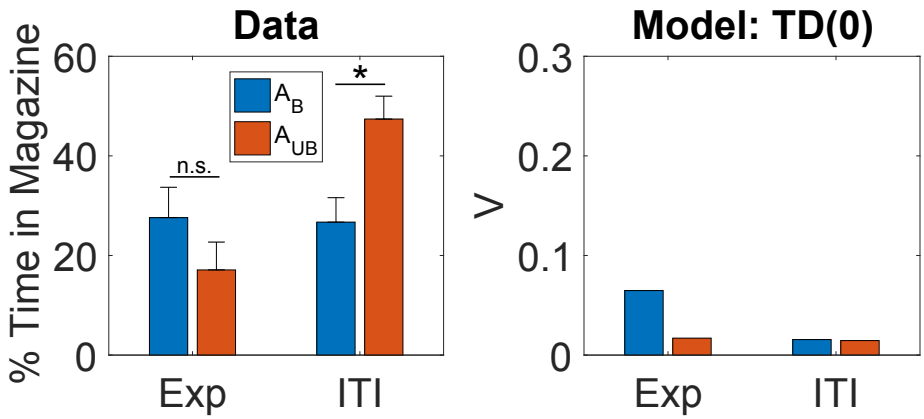

Figure 1: **Inhibition of dopamine neurons prevents learning induced by changes in reward identity**. (*Left*) Conditioned responding on the probe test in the identity unblocking paradigm. Exp: experimental group, receiving inhibition during reward outcome. ITI: control group, receiving inhibition during the intertrial interval. Asterisk indicates significant difference ($p < 0.05$). Data replotted from [1]. (*Right*) Model simulation of the value function.

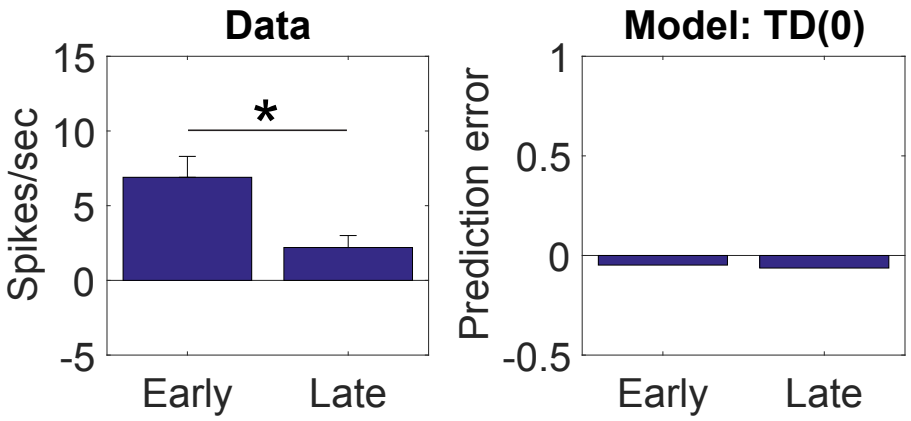

Figure 2: **Dopamine neurons respond to changes in reward identity**. (*Left*) Firing rate of dopamine neurons on trials that occurred early (first 5 trials) or late (last 5 trials) during an identity shift block. Data replotted from [2]. (*Right*) Model simulation of TD error.

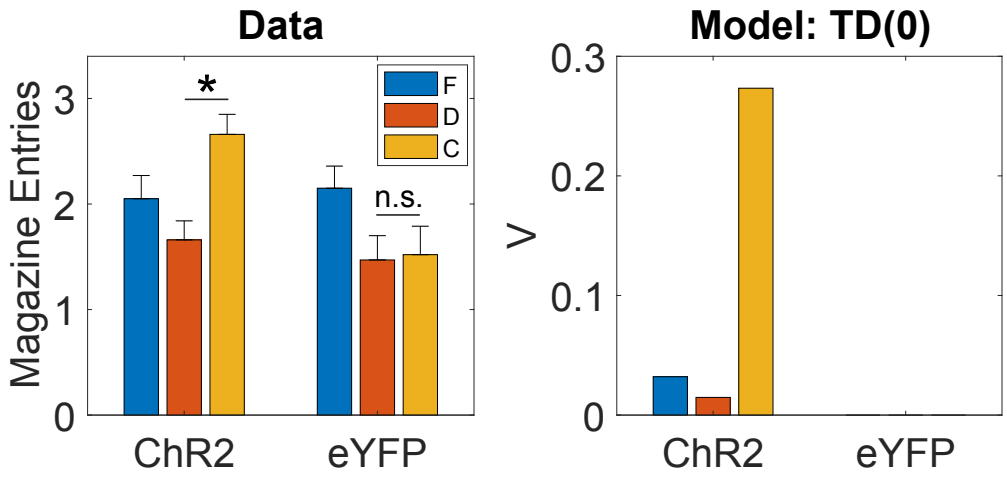

Figure 3: **Dopamine transients are sufficient for learning stimulus-stimulus associations**. (*Left*) Number of food cup entries occurring during the probe test for experimental (ChR2) and control (eYFP) groups in the sensory preconditioning paradigm. Data replotted from [3]. (*RIght*) Model simulation, using the value estimate as a proxy for conditioned responding. Note that V attached to the critical cue, C, is high in the simulation, much like the food cup responding in the probe test to this cue. This occurs because dopamine is paired with the cue, so it directly acquires a significant value. However, in this paradigm there is no direct link between C and the policy of going to the food cup. Thus, the success of TD(0) in this context in matching the empirical data is somewhat misleading.

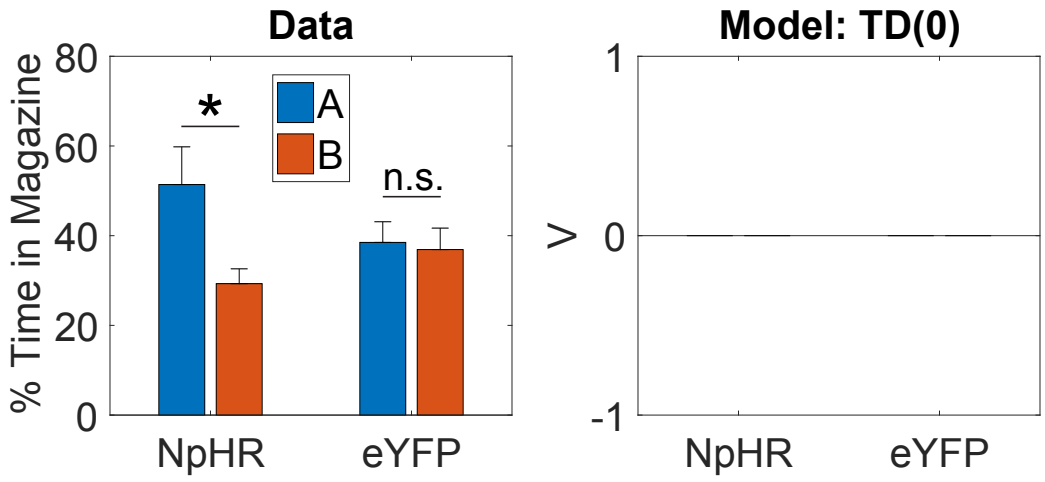

Figure 4: **Dopamine transients are necessary for learning stimulus-stimulus associations**. (*Left*) Number of food cup entries occurring during the probe test for experimental (NpHR) and control (eYFP) groups in the sensory preconditioning paradigm. Data replotted from [3]. (*Right*) Model simulation.

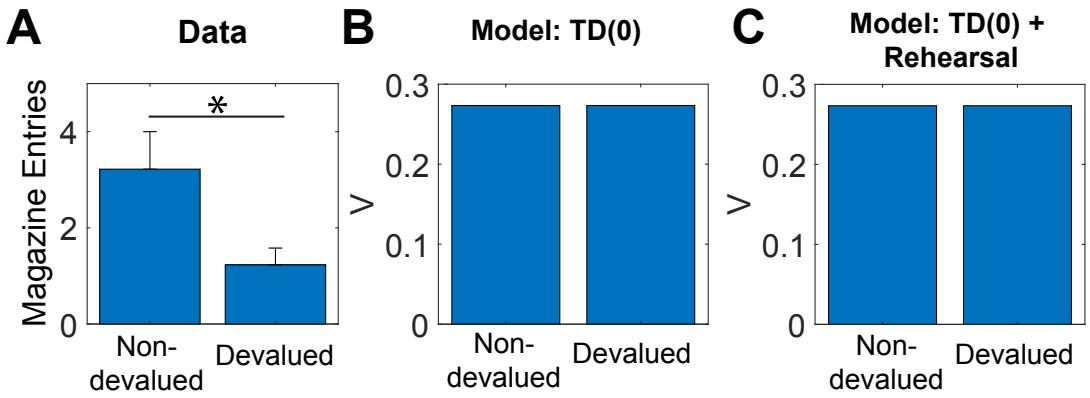

Figure 5: **Behavior to preconditioned cue that is unblocked by activation of dopamine neurons is sensitive to devaluation of the predicted reward**. Data (A, replotted from [3]) and model simulation (B) for conditioned responding to stimulus C in the probe test. Animals in the devalued group were injected with lithium chloride in conjunction with ingestion of the reward (sucrose pellets), causing a strong aversion to the reward. Animals in the nondevalued group were injected with lithium chloride approximately 6 hours after ingestion of the reward. (C) A version of the model with rehearsal of stimulus X during reward devaluation was able to capture the devaluation-sensitivity of animals.

References

[1] Chun Yun Chang, Matthew Gardner, Maria Gonzalez Di Tillio, and Geoffrey Schoenbaum. Optogenetic blockade of dopamine transients prevents learning induced by changes in reward features. *Current Biology*, 27:3480–3486, 2017.

[2] Yuji K Takahashi, Hannah M Batchelor, Bing Liu, Akash Khanna, Marisela Morales, and Geoffrey Schoenbaum. Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. *Neuron*, 95:1395–1405, 2017.

[3] Melissa J Sharpe, Chun Yun Chang, Melissa A Liu, Hannah M Batchelor, Lauren E Mueller, Joshua L Jones, Yael Niv, and Geoffrey Schoenbaum. Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience*, 20, 2017.