Title

Spatial localization of hippocampal replay requires dopamine signaling

Authors

Matthew R. Kleinman[1] and David J. Foster[1,2]

Affiliations

1 Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley, CA 94720, USA

2 Lead contact

Correspondence

davidfoster@berkeley.edu, mattrkleinman@berkeley.edu

Abstract

Sequenced reactivations of hippocampal neurons called replays, concomitant with sharp-wave ripples in the local field potential, are critical for the consolidation of episodic memory, but whether replays depend on the brain's reward or novelty signals is unknown. Here we combined chemogenetic silencing of dopamine neurons in ventral tegmental area (VTA) and simultaneous electrophysiological recordings in dorsal hippocampal CA1, in freely behaving rats experiencing changes to reward magnitude and environmental novelty. Surprisingly, VTA silencing did not prevent ripple increases where reward was increased, but caused dramatic, aberrant ripple increases where reward was unchanged. These increases were associated with increased reverse-ordered replays. On familiar tracks this effect disappeared, and ripples tracked reward prediction error, indicating that non-VTA reward signals were sufficient to direct replay. Our results reveal a novel dependence of hippocampal replay on dopamine, and a role for a VTA-independent reward prediction error signal that is reliable only in familiar environments.

Introduction

Spatial information is encoded in the firing of hippocampal place cells, which are thought to provide a cognitive map to support memory and navigation (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978). During pauses in locomotion, place cells participate in structured population bursts of activity, representing temporally-compressed trajectories through experienced locations, a phenomenon termed replay (Diba and Buzsáki, 2007; Foster and Wilson, 2006). Replay sequences can occur in the same order as experience, called forward replay, or in the reverse order of experience, called reverse replay (Diba and Buzsáki, 2007; Foster and Wilson, 2006). Replay appears after just one experience in a novel environment (Berners-Lee et al., 2022), and is preferentially generated towards goals in goal-directed tasks (Pfeiffer and Foster, 2013; Widloski and Foster, 2022). Experimentally interrupting or lengthening replay-associated sharp wave ripples (SWR) in the local field potential (LFP) disrupts and enhances learning of a spatial memory task, respectively (Fernández-Ruiz et al., 2019; Jadhav et al., 2012). Replay is thus thought to support memory consolidation and online planning (Buzsáki, 2015; Foster, 2017).

Intriguingly, reward drives increased rates of SWR (Singer and Frank, 2009), and only reverse replay, not forward, is increased at highly rewarding locations (Ambrose et al., 2016). Theoretical work suggests replay functions to update spatial representations of value to influence behavior and optimize reward receipt (Mattar and Daw, 2018). These findings hint that replay may be strongly modulated by reward-processing areas in the brain, such as the midbrain dopamine system (Fields et al., 2007). Dopamine neuron activity in the ventral tegmental area (VTA) is consistent with coding of reward prediction error (RPE), with increased activity at unexpected rewards and decreased activity with omission of expected rewards (Schultz et al., 1997). Subsequent work investigated dopamine release in spatial tasks, finding it ramps towards large rewards (Guru et al., 2020; Howe et al., 2013) in a manner consistent with encoding of RPE for a value function over space (Kim et al., 2020).

Besides the well-established role of midbrain dopamine neurons in reward processing, dopamine release in hippocampus has been implicated in stabilizing place fields (Kentros et al., 2004), gating the increase in plasticity in dorsal CA1 synapses by novel experiences (Li et al., 2003), and improving memory retention via increasing replay (McNamara et al., 2014). Furthermore, VTA activity is increased in novel environments (Guru et al., 2020; McNamara et al., 2014; Takeuchi et al., 2016), suggesting the hippocampus and VTA may coordinate to signal spatial novelty and induce learning in new environments (Lisman and Grace, 2005). However, recent work implicates locus coeruleus (LC) as the dominant source of dopaminergic input to dorsal CA1 and show its necessity and sufficiency for novelty-mediated episodic memory consolidation (Guru et al., 2020; McNamara et al., 2014; Takeuchi et al., 2016), leaving the role of VTA unclear.

We therefore tested whether VTA dopamine neurons are required for reward-related modulation of SWR and replay. We expressed an inhibitory DREADD (Armbruster et al., 2007) in VTA dopamine neurons and implanted a tetrode microdrive above hippocampus. We could then inhibit VTA dopamine signaling and simultaneously record neural activity in the dorsal CA1 region of hippocampus while rats collected rewards in familiar and novel environments. If VTA dopamine signaling is required for coordinating replay to valuable locations, we expected to see deficits in the capacity for reward to recruit SWR and replay. Additionally, if VTA dopamine is critical for inducing plasticity in CA1 in novel environments, novelty may significantly increase the effect of VTA inactivation on hippocampal replay.

Results

VTA inactivation in a simple spatial task with reward changes

We combined tetrode recordings in dorsal CA1 (dCA1) and chemogenetic silencing of VTA dopamine neurons to determine whether reward-related changes in hippocampal ripples and replay required VTA dopamine signaling. Transgenic rats expressing cre-recombinase under the tyrosine hydroxylase (TH) promoter were stereotactically injected with cre-dependent virus containing the inhibitory DREADD hM4Di (Experimental, n=4) or mCherry-only control (Control, n=3) into bilateral VTA (Figure 1A). We observed widespread expression across the extent of VTA and co-localization with TH (Figure 1B), enabling specific and reversible inactivation of VTA dopamine signaling. Recording microdrives containing 6-32 independently adjustable

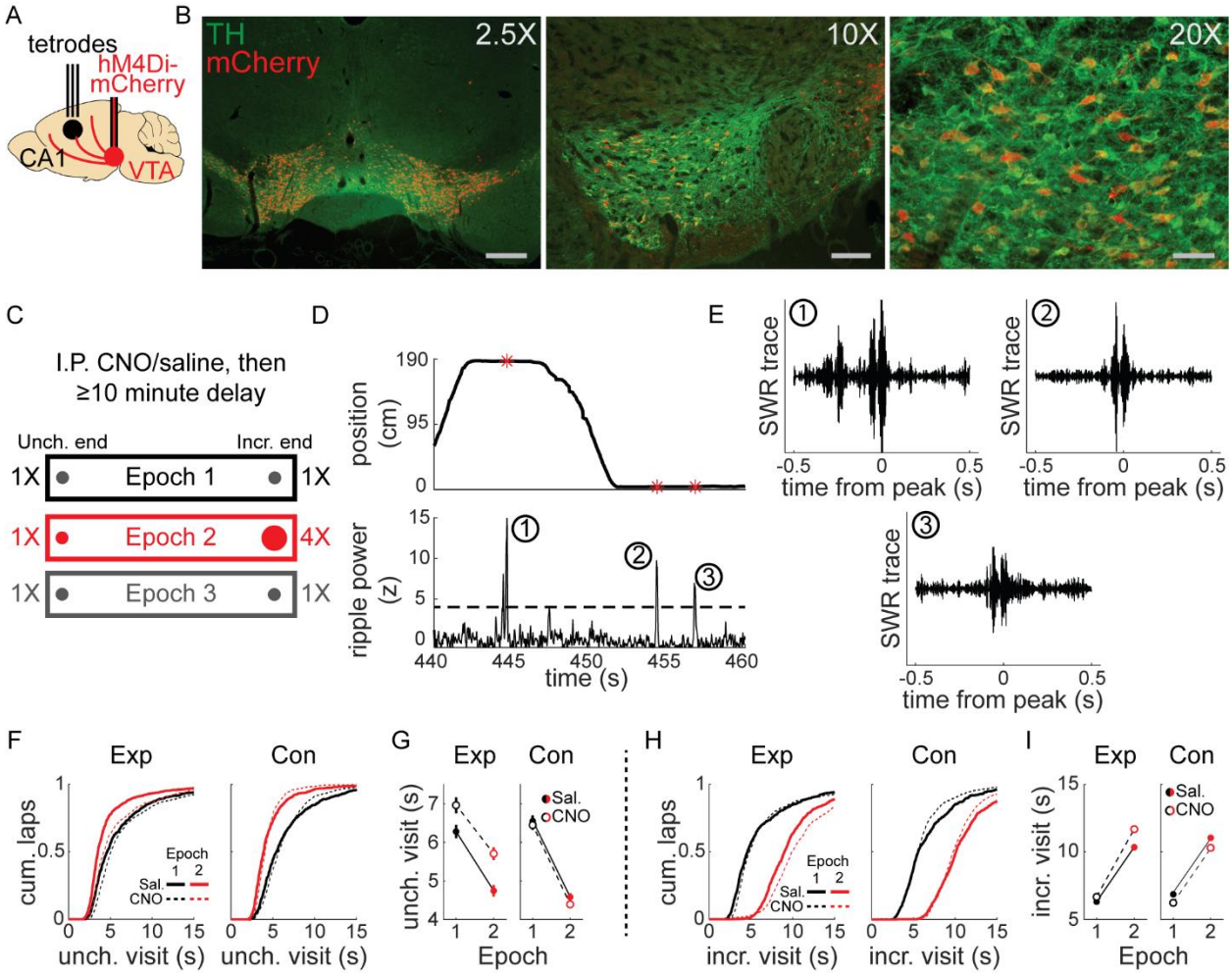


Figure 1. Experimental design and linear track behavior.

(A) TH-cre rats underwent stereotactic surgery to inject virus bilaterally into VTA and implant a tetrode microdrive above dorsal CA1. **(B)** Co-expression of mCherry (red) and TH (green) in VTA from three example animals. Left panel, mCherry-only virus, scale bar 600 µm; middle panel, hM4Di-mCherry, scale bar 150 µm; right panel, hM4Di-mCherry, scale bar 75 µm. **(C)** Intraperitoneal injection of saline or CNO (1-4 mg/kg) preceded recording sessions by at least 10 minutes. Rats were placed at one end of a linear track and collected liquid chocolate reward from wells at each end. Each epoch lasted 10-20 laps and reward changes were unsignaled to the animal. For each session, the Incr. end was defined as the reward end with 4X reward in Epoch 2, and the Unch. end was defined as the reward end with 1X reward in Epoch 2. **(D)** During stopping periods at reward ends, LFP was bandpass filtered in the ripple band (150-250 hz) and SWR events were detected. **(E)** Three example ripple-filtered LFP traces from one lap (two stopping periods) are shown. **(F)** Cumulative distribution of reward end stopping periods at the Unch. reward end in Epoch 1 and 2 for experimental rats (left panel) and control rats (right panel). See also Figures S1-S3. **(G)** The duration of Unch. reward end stopping periods decreased from Epoch 1 to Epoch 2. Mean ± standard error, Exp Saline, Epoch 1: 6.28±0.17, Epoch 2: 4.74±0.15, two-sample t-test: $t(1530)=6.7$, $p<10^{-8}$; Exp CNO, Epoch 1: 6.96±0.21, Epoch 2: 5.71±0.16, two-sample t-test: $t(1352)=4.785$, $p<10^{-5}$. Con Saline, Epoch 1: 6.55±0.15, Epoch 2: 4.58±0.1, two-sample t-test: $t(1149)=11.032$, $p<10^{-10}$; Con CNO, Epoch 1: 6.45±0.12, Epoch 2: 4.39±0.06, two-sample t-test: $t(1286)=15.06$, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch ($F[1,5317]=252.26$, $p<10^{-10}$), drug ($F[1,5317]=9.93$, $p=0.0016$), group ($F[1,5317]=16.09$, $p=0.0001$), epoch X group ($F[1,5317]=8.23$, $p=0.0041$), drug X group ($F[1,5317]=20.3$, $p<10^{-5}$). **(H)** Cumulative distribution of reward end stopping periods at the Incr. reward end in Epoch 1 and 2 for experimental rats (left panel)

and control rats (right panel). **(I)** The duration of Incr. reward end stopping periods increased from Epoch 1 to Epoch 2. Mean ± standard error, Exp Saline, Epoch 1: 6.314±0.17, Epoch 2: 10.351±0.2, two-sample t-test: $t(1514)=-15.315$, $p<10^{-10}$; Exp CNO, Epoch 1: 6.67±0.22, Epoch 2: 11.691±0.25, two-sample t-test: $t(1340)=-15.059$, $p<10^{-10}$. Con Saline, Epoch 1: 6.859±0.17, Epoch 2: 11.047±0.17, two-sample t-test: $t(1138)=-17.447$, $p<10^{-10}$; Con CNO, Epoch 1: 6.229±0.12, Epoch 2: 10.304±0.11, two-sample t-test: $t(1274)=-24.745$, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch ($F[1,5266]=1077.4$, $p<10^{-10}$), drug X group ($F[1,5266]=33.8$, $p<10^{-5}$), epoch X drug X group ($F[1,5266]=4.33$, $p=0.0376$).

tetrodes (bilateral 32 tetrode, n=4; unilateral 20 tetrode, n=2; unilateral 6 tetrode, n=1) were implanted above dCA1 (Fig. 1A). Each tetrode was lowered into the pyramidal cell layer of dCA1 to collect single unit and LFP data.

Before each experimental session, rats were given intraperitoneal injection of CNO, to activate hM4Di receptors and suppress VTA dopamine neuron activity, or saline, then performed a simple task on linear tracks (1.5 to 2.5 m in length), collecting liquid chocolate rewards from each end (Fig. 1C). Each session began with equal 0.1 ml reward volume at each end (Epoch 1) for 10-20 laps (1 lap was comprised of reward collection at both ends; mean, 16 laps). This was followed by unsignaled quadrupling of reward at one end to 0.4 ml (Incr. end), while reward at the other end remained unchanged (Unch. end), for 10-20 laps (Epoch2; mean, 16.7 laps). Finally, reward was equalized again to 0.1 ml at both ends (Epoch 3) for up to 20 laps (mean, 11.6 laps).

Each animal performed this task on familiar linear tracks (>2 sessions on track; total session count for each condition: Experimental rats: 36 saline, 34 CNO; Control rats: 23 saline, 25 CNO) and novel linear tracks (1st or 2nd session on track; Experimental rats: 9 saline, 10 CNO; Control rats: 12 saline, 12 CNO). During stopping periods at either end of the track (velocity ≤ 8 cm/s, position ≤ 10 cm from end), SWR were identified as peaks in the ripple band (150-250 hz) in LFP (Fig. 1D-E; see Methods).

Gross behavior was largely unaffected by VTA suppression (e.g., all reward consumed on each lap), but CNO in experimental animals systematically affected stopping period duration. Visits to the Unch. end in Epoch 2 were significantly shorter than in Epoch 1, despite unchanged reward volume there, and while this reduction was present in CNO sessions, overall visit durations were increased (Fig. 1F-G). CNO did not affect control animals (Fig. 1F-G). Visits to the Incr. end were significantly longer in Epoch 2 than in Epoch 1 in all conditions, owing to the increased reward consumption time (Fig. 1H-I; Epoch 1 vs. Epoch 2, two sample t-test, all $p<10^{-10}$). Changes in stopping period duration in Epoch 3 were similar across all conditions: Unch. visit duration increased from Epoch 2 to Epoch 3 (Fig. S1E), while Incr. visit duration decreased (Fig. S1F). Separate analysis of novel and familiar sessions revealed the pattern of shorter duration Unch. visits in Epoch 2 compared to Epoch 1 did not depend on novelty (Fig. S1A-D). However, the main effect of CNO in experimental rats of prolonging stopping periods occurred in novel sessions (Fig. S1A-B), not in familiar sessions (Fig. S1C-D). Rats consistently ran slightly faster towards the Incr. end than the Unch. end in Epoch 2, across all conditions (Fig. S2).

We interpret the reduction in visit duration as a behavioral signature of the Unch. end becoming relatively less valuable during Epoch 2, when the reward volume is larger at the Incr. end. Though visit durations were slightly longer in CNO sessions in

experimental rats, particularly in novel track sessions, this behavioral effect of a relative value decrease remained, indicating VTA inactivation did not prevent rats from recognizing a devalued location.

Reward-related modulation of SWR rate is mediated by novelty and VTA

We analyzed the rate of SWR occurrence during the first 10 s of each stopping period, when rats were consuming reward. In individual sessions, SWR rate increased robustly in all conditions shortly after stopping at the reward wells and beginning reward consumption (Fig. 2A). Relative to Epoch 1, SWR rate during Epoch 2 tended to increase dramatically at Incr. end visits and decrease at Unch. end visits. During Epoch 3, SWR rate at the Incr. end dropped precipitously relative to Epoch 2, while often increasing at the Unch. end (Fig. S4A).

Surprisingly, during novel sessions, VTA inactivation often led to increased SWR rate at both reward ends (Fig. 2A, right). SWR rate still increased in Epoch 2 at the Incr. end even without normal VTA signaling, indicating reward sensitivity per se was not abolished, but suggesting the localization of this increased SWR rate to where reward increased was disrupted.

Pooling across sessions revealed this dramatic increase in SWR rate at the Unch. end was typical with CNO in novel sessions, and further suggested there was a reduction in the difference in SWR rate between the Incr. and Unch. ends in Epoch 2 in both familiar and novel experiences (Fig. 2B). We therefore used a Poisson generalized linear model (GLM) to quantify the changes in SWR rate across reward end, epoch, drug condition, and novelty (see Methods). In experimental rats, CNO and reward both influenced SWR rate, with significant effects for the CNO main effect ($z=3.19$, $p=0.0014$), the interaction between Incr. end and Epoch 2 ($z=9.02$, $p<10^{-10}$), the three-way interaction between Incr. end, Epoch 2, and CNO ($z=-2.06$, $p=0.0396$), and marginally by the interaction between Incr. end and CNO ($z=-1.92$, $p=0.055$). Control animals showed no apparent effect of CNO (Fig. 2C). The same Poisson GLM fit to control rat data confirmed this, with significant coefficients only for Incr. end ($z=-2.42$, $p=0.0156$) and the interaction between Incr. end and Epoch 2 ($z=7.64$, $p<10^{-10}$).

To assess the interaction between VTA inactivation and novelty, we fit the Poisson GLM separately to novel and familiar sessions, then used bootstrapping to generate distributions of SWR rates for each reward end and condition under the null hypothesis that CNO had no effect (see Methods). We found the actual difference in SWR rates between saline and CNO sessions in experimental animals during epoch 2 was significantly greater than chance (Fig. S3A) in novel sessions for both the Unch. end (one-tailed test, CNO>saline, $p=0.0006$) and the Incr. end (one-tailed test, saline>CNO, $p=0.004$), as well as at the Unch. end in familiar sessions (one-tailed test, CNO>saline, $p=0.002$). There was no significant difference between saline and CNO in control animals at either reward end in either familiar or novel sessions (Fig. S3B; one-tailed tests, all $p>0.17$).

A potential functional role for the reward-related changes in SWR rate is to strengthen downstream representations of particularly rewarding locations at the expense of less rewarding locations. We tested whether VTA suppression blunted the SWR rate difference between reward ends in Epoch 2. Across both familiar and novel

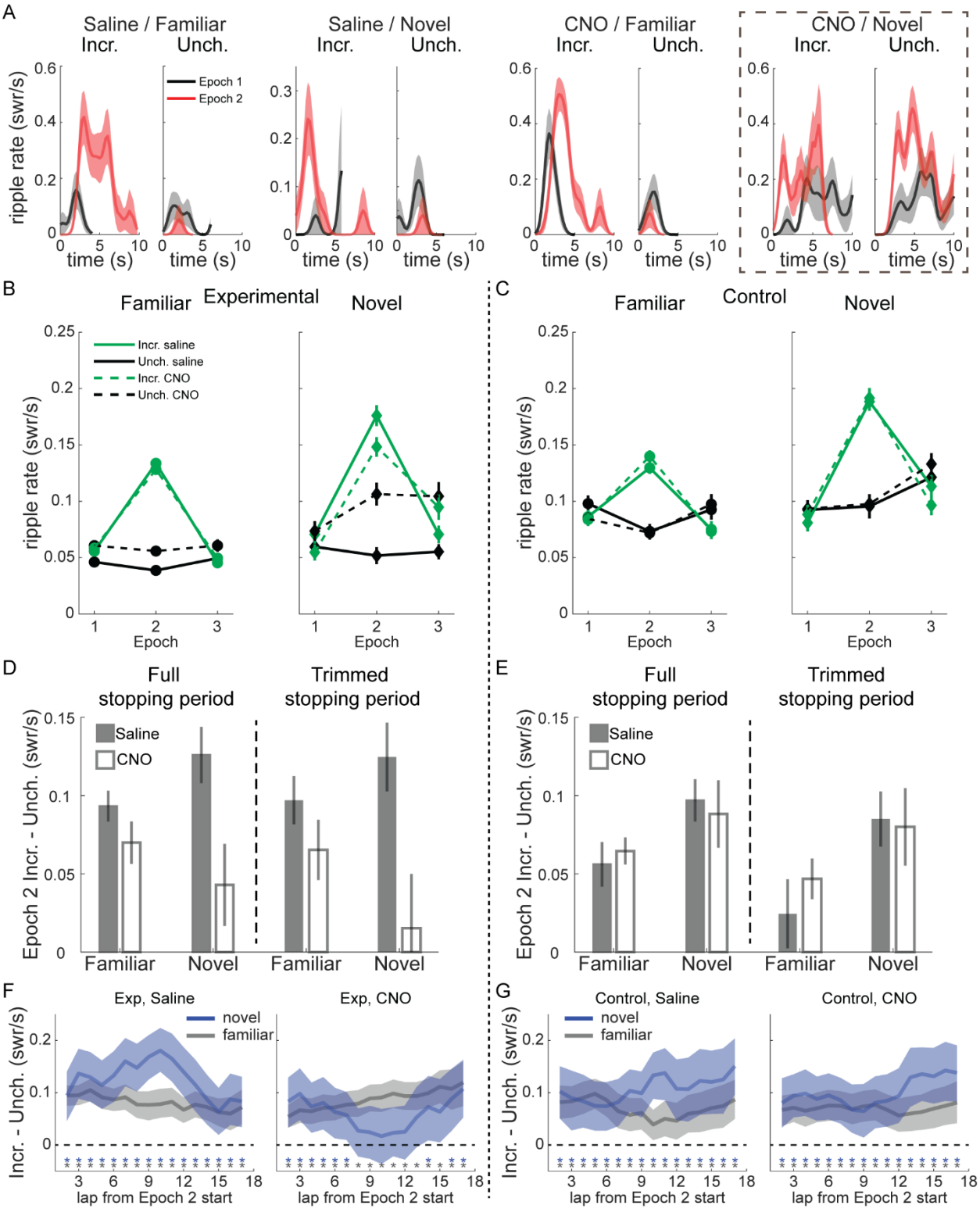


Figure 2. Modulation of SWR rate by reward, novelty, and VTA inactivation.
(A) SWR rate as a function of time in stopping period in Epoch 1 and 2 for four example sessions in experimental rats; from left to right, saline on familiar track, saline on novel track, CNO on familiar track, and CNO on novel track. In each panel, visits to the Incr. end are on the left and visits to the Unch. end are on the right. Relative to Epoch 1 (black lines), in Epoch 2 (red lines) SWR rate increased at Incr. end

and decreased at Unch. end in all conditions except for CNO on a novel track (far right), where SWR rate increased at both ends in Epoch 2. SWR rate was binned in 0.25 s windows and smoothed with a two-bin Gaussian. Line, mean; shading, standard error. **(B)** SWR rate in experimental rats as a function of epoch, drug (saline in solid lines, CNO in dashed lines), reward end (Unch. in black, Incr. in green), and novelty (familiar in left panel, novel in right panel). See also Figure S3 and S4. **(C)** SWR rate in control rats as a function of epoch, drug (saline in solid lines, CNO in dashed lines), reward end (Unch. in black, Incr. in green), and novelty (familiar in left panel, novel in right panel). **(D)** Difference between SWR rate at Incr. and Unch. ends in Epoch 2 in Experimental rats. Full stopping period, left panel. Trimmed stopping period, with first 1 s and last 1 s of visit excluded to eliminate all slow approaching/leaving movement, right panel. Saline, gray bars; CNO, white bars. Mean and standard error. Full stopping periods, three-way ANOVA with animal group, drug, and novelty: drug ($F[1,153]=5.19$, $p=0.0241$), group X drug ($F[1,153]=5.16$, $p=0.0245$). Trimmed stopping periods, three-way ANOVA with animal group, drug, and novelty: group X drug ($F[1,153]=5.58$, $p=0.0194$). **(E)** Difference between SWR rate at Incr. and Unch. ends in Epoch 2 in Control rats, as in (D). Statistics in legend (D). **(F)** In experimental rats, the difference in SWR rates at each reward end (Incr. – Unch.) in Epoch 2, after subtracting the mean rates in Epoch 1, averaged over a 5-lap sliding window within Epoch 2. Blue lines, novel sessions. Gray lines, familiar sessions. Blue and gray asterisks denote the centers of sliding windows in which the difference in SWR rate was significantly greater than 0 in novel and familiar sessions, respectively (one-sample t-test, $p<0.05$). Shading denotes 95% confidence interval. See also Figure S4. **(G)** As in (F), but for control animals.

environments, CNO reduced the difference in SWR rate between the Incr. end and Unch. end in experimental rats but not in control rats (Fig. 2D-E, left panels; three-way ANOVA with animal group, drug, and novelty: drug, $F[1,153]=5.19$, $p=0.024$, group by drug, $F[1,153]=5.16$, $p=0.025$, all other terms n.s.). To control for the possibility that VTA inactivation caused changes in locomotor or other non-consummatory behavior at reward wells that might affect SWR emission, we omitted the first and last 1 s of each stopping period to isolate the reward consumption period. The effect of CNO remained in experimental rats, reducing SWR rate discrimination between Incr. and Unch. ends (Fig. 2D-E, right panels; three-way ANOVA with animal group, drug, and novelty: drug, $F[1,153]=3.41$, $p=0.067$; group by drug, $F[1,153]=5.58$, $p=0.019$).

We next looked for within-epoch changes in SWR rate to determine whether VTA inactivation altered the dynamics of the response to reward changes. We calculated the difference in SWR rate at the Incr. and Unch. reward ends (each with its Epoch 1 mean subtracted) in a 5-lap sliding window. In all time windows and conditions except novel CNO sessions in experimental rats, the SWR rate at the Incr. end was significantly greater than at the Unch. end (Fig. 2F-G; Incr. – Unch. significantly greater than 0, one-sample t-test, $p<0.05$, uncorrected for multiple comparisons).

In novel sessions, VTA inactivation did not prevent an initially larger increase in SWR rate at the Incr. end than the Unch. end, but caused that difference to diminish over laps (Fig. 2F). By the middle of the epoch, there was no statistically-significant difference in reward modulation between the reward ends (one-sample t-test, $p>0.05$, for 5-lap windows centered on laps 8-13 and 15), consistent with an initial appropriately-localized reaction to reward change that eventually spread across the track. We found a similar deficit in Epoch 3, with SWR rate decreasing significantly more compared to Epoch 2 at the Incr. end than the Unch. end (Incr. – Unch. significantly below 0, one-sample t-test, $p<0.05$) for almost every task condition and timepoint except in novel CNO sessions in experimental rats (Fig. S4B-C). This suggests VTA inactivation may

disrupt the normal magnitude or time course of the SWR response to negative value changes as well.

Overall, VTA inactivation spared the capacity for increased reward to modulate SWR rate but led to decreased differentiation of low and high value locations, particularly in novel environments where SWR rate increased spatially-indiscriminately.

SWR rate is correlated with RPE even with VTA inactivation

Taken together, the above results demonstrate VTA inactivation caused changes in the normal dynamics of the response of SWR rate to positive and negative changes in reward value (Fig. 2). However, because each session had only two timepoints when reward value changed by fixed amounts, our experiment was not optimized to probe the precise relationship between SWR rate and reward changes. Additionally, the effect of VTA inactivation was particularly prominent with novelty, when both SWR rate and its modulation by reward changes were greater, raising the possibility that large SWR rates and fluctuations, rather than novelty per se, depend on VTA dopamine signaling.

To address these questions, we designed a volatile reward schedule (Experiment 2) with frequent, large reward changes at one end of the linear track, and tested whether VTA inactivation impacted the capacity for SWR rate to track value (Fig. 3A, top). The “stable end” delivered 0.2 ml every lap, while the “volatile end” reward volume was drawn pseudorandomly from 0, 0.1, 0.2, 0.4, and 0.8 ml (mean 0.37 ml; blocks of 20 laps were comprised of 3 laps x 0 ml, 4 laps x 0.1 ml, 3 laps x 0.2 ml, 4 laps x 0.4 ml, and 6 laps x 0.8 ml). The position of the stable and volatile ends randomly varied across sessions.

This reward schedule also allowed us to test whether SWR rate was correlated with value, RPE, or neither. We expected SWR rate at the volatile end would be predominantly determined by the current reward volume there, but potentially also modulated by previous reward volumes (Fig. 3A, bottom). If SWR rate is correlated with value, then for a given current volume, larger reward volumes at the last visit will lead to higher SWR rates compared to when the last visit was a smaller reward volume. Conversely, if SWR rate is correlated with RPE, the opposite modulation by last reward volume will be observed: the larger the previous reward volume, the lower the current SWR rate.

A subset of rats performed sessions of the modified reward schedule (mean 53.4 laps per session; Total sessions per condition: Experimental rats 2 and 4: 6 saline, 7 CNO; Control rats 1-3: 16 saline, 18 CNO). Each rat completed 1-2 saline sessions before any CNO sessions and all sessions were on the same linear track, meaning almost all CNO sessions took place on a familiar track. As expected, SWR rate at the volatile end was predominantly determined by the current volume (Fig. 3B-C). There was little obvious difference between saline and CNO in either experimental (Fig. 3B) or control rats (Fig. 3C). SWR rate at the stable end was largely stable across laps, although there was a trend towards higher SWR rate if the most recent volatile end visit was lower volume, consistent with lap-by-lap changes in the relative value of the stable end (Fig. S5).

We next investigated how SWR rate at the volatile end varied as a function of both current and immediately previous volatile end volume in experimental rats (Fig. 3D-E, top panels). For each current volume, we subtracted the mean SWR rate across all

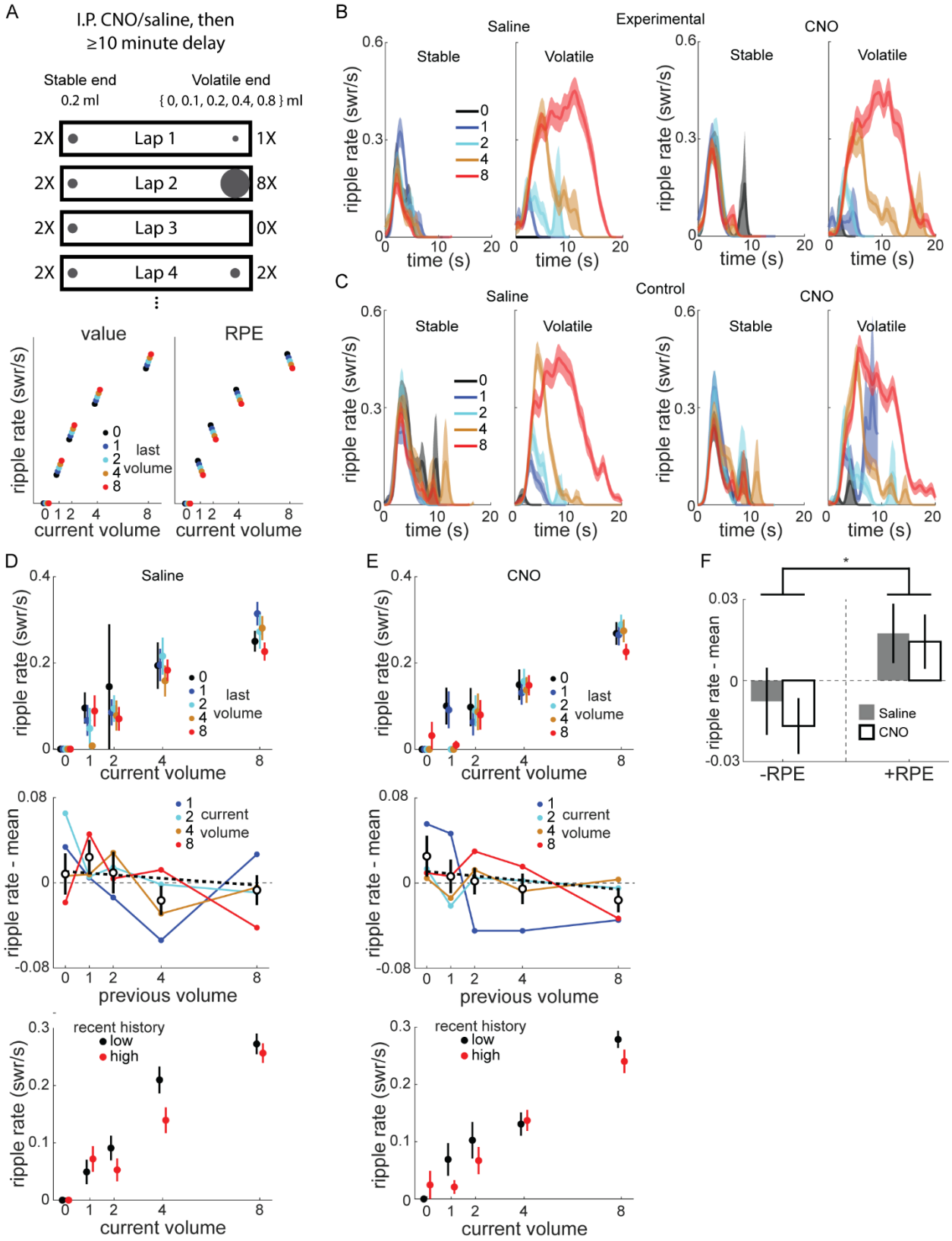


Figure 3. Frequent reward changes modulated SWR rate.

(A) Recording sessions in the volatile reward task were preceded by intraperitoneal injection of saline or CNO by at least 10 minutes. Rats were placed on the stable end to begin each session, which delivered 0.2 ml reward at each visit, while the volatile end delivered 0, 0.1, 0.2, 0.4, or 0.8 ml, pseudorandomly chosen on each lap. Bottom panel, schematic of how value and RPE would modulate SWR. Given a particular current volume, value coding predicts a positive correlation between SWR rate and previous volume, while RPE coding predicts a negative correlation. **(B)** SWR rate as a function of reward volume and time in end visit in example rat, experimental rat 4. Left panel, saline. Right panel, CNO. In stable panel, traces are colored based on previous volatile end visit volume. In volatile panel, traces are colored based on current volatile volume. See also Figure S5 and S6. **(C)** SWR rate as a function of reward volume and time in end visit in example control rat 3, as in (B). **(D)** Top panel, SWR rate at volatile end as a function of current and previous volatile volume, for saline sessions in experimental rats. Middle panel, SWR rate for each non-zero volatile volume plotted as a function of previous volume, with the mean SWR rate for that current volume subtracted. Unfilled symbols, mean of previous volume across all current volumes. Thick dashed line, linear fit to mean values. Pearson correlation between (ripple rate – mean) and previous volume, r=-0.076, p=0.177. Error bars, standard error. Bottom panel, SWR rate as a function of reward volume, separated by recent reward history (median split on average of last 3 visits). Black, recent history below median; red, recent history above median. **(E)** Same as (D), for CNO sessions in experimental rats. Middle panel, Pearson correlation between (ripple rate – mean) and previous volume, r=-0.109, p=0.049. GLM fitting SWR rate as a function of drug, current volume, and previous volume: previous volume, z=-2.31, p=0.021; drug and current volume, both p>0.8. Bottom panel, Poisson GLM fitting ripple rate as a function of volume, drug condition, and reward history (above/below median): volume, z=13.86, $p<10^{-10}$; history, z=-2.23, p=0.026; drug, z=-1.05, p=0.29. **(F)** The RPE of volatile end visits were calculated by subtracting the previous volatile volume from the current volume. Two-way ANOVA with drug and RPE sign (+/-): drug (F[1,518]=0.3, p=0.582), RPE sign (F[1,518]=6.42, p=0.0116), drug X RPE sign (F[1,518]=0.07, p=0.785).

previous volumes, and examined how previous volume affected the mean-subtracted SWR rates across all current volumes (Fig. 3D-E, middle panels).

The mean-subtracted SWR rate was modestly negatively correlated with the previous volume (Pearson correlation: saline, r=-0.076, p=0.177; CNO, r=-0.109, p=0.049). A GLM found the mean-subtracted SWR rate was significantly affected by previous volume (z=-2.3, p=0.02), but not drug (z=-0.24, p=0.81) or current volume (z=0.196, p=0.844). There was no similar behavioral effect: for each current volume, we subtracted the mean reward end visit duration, and found no correlation between the previous reward volume and mean-subtracted visit duration (Pearson correlation; saline, r=0.011, p=0.844; CNO, r=-0.075, p=0.175).

We next separated visits to the volatile end using a median split based on the recent volumes (mean of previous 3 visits). SWR rates were higher for a given current reward volume when the recent reward history was low (Fig. 3D-E, bottom panels). A Poisson GLM predicting SWR rate as a function of current volume, drug condition, and whether reward history was low or high (relative to the median for the session) revealed significant effects of current volume (z=13.86, $p<10^{-10}$), as expected, and reward history split (z=-2.23, p=0.026), but not drug condition (z=-1.05, p=0.29).

Finally, we separated combinations of current and previous volume into those with negative RPE (current < previous) and positive RPE (current > previous) and found mean-subtracted SWR rate was significantly affected by RPE sign (Fig. 3F; two-way ANOVA with drug and RPE sign, RPE sign: F[1,518]=6.42, p=0.012), but not drug (F[1,518]=0.3, p=0.582; drug X RPE sign, F[1,518]=0.07, p=0.785).

Given the lack of an effect of drug in experimental rats, we pooled all sessions (both animal groups, both drug conditions) in Experiment 2 to maximize experimental

power and found similar results as in just experimental rats (Fig. S6). On top of a large increase in SWR rate with current volume at the volatile end (Fig. S6A-B), SWR rate was also significantly negatively correlated with the previous volatile volume, both at the volatile end (Fig. S6C) and at the stable end (Fig. S6E). Accordingly, SWR rates were significantly lower for negative RPE than for positive/non-negative RPE (Fig. S6D-F). Finally, recent reward history at the volatile end significantly affected SWR rate (Fig. S6G), with higher SWR rate when recent rewards were lower.

Taken together, SWR rate was modulated by reward volume changes consistent with RPE-like coding. This modulation did not require normal VTA dopamine signaling, at least in familiar environments. The lack of effect of VTA inactivation, even with frequent, large swings in value and SWR rate, corroborates the results from Experiment 1 that novel experiences are particularly susceptible to disruption, indicating VTA dopamine release is critical when learning new reward locations.

Rate of reverse replay is increased with reward in novel environments only with intact VTA signaling

Previous work discovered the incidence rate of reverse replay, but not forward replay, was increased at locations with increased reward (Ambrose et al., 2016). We therefore analyzed single unit data collected in Experiment 1 (excluding experimental rat 2, who had a 6-tetrode recording drive) to determine whether this modulation of replay required VTA dopamine signaling. As previously observed (e.g., Ambrose et al., 2016), place cells in dCA1 had directional fields, such that the location a neuron was active while the rat moved in one direction on the track (e.g., “upward”) was often distinct from its activity when the rat moved in the other direction (e.g., “downward”). This directionality was apparent in both familiar and novel sessions, including in experimental rats with either saline or CNO (Fig. 4A). We found no effect of CNO on within-session field reliability, but significantly less reliability in novel compared to familiar sessions (Fig. S7A). Field similarity across running directions was slightly but significantly increased by both CNO and novelty (Fig. S7B).

Two place fields were defined for each neuron, one for each running direction, permitting Bayesian decoding methods to estimate both position and direction from neural activity (Fig. 4B-C). Sessions with accurate position and direction decoding during run, primarily due to sufficiently high cell yield, were included for replay analysis (Fig. S8; total sessions included: Experimental rats: novel saline, n=8; novel CNO, n=8; familiar saline, n=18; familiar CNO, n=23; Control rats: novel saline, n=12; novel CNO, n=11; familiar saline, n=16; familiar CNO, n=17).

Candidate replay events were periods of high population spiking activity while the rat was not running (z-scored spike count>3, minimum duration of 50 ms, rat velocity≤8 cm/s). We used a memory-less Bayesian decoder, with 40 ms decoding windows advancing by 5 ms steps, to estimate position and direction from neural activity. Replays were defined as candidate events with spatial trajectories meeting a threshold for motion and minimum total posterior in one running direction map 13 (see Methods), with the running direction with greater posterior probability used to classify replay directionality, described below.

Forward replays were spatial trajectories moving across the track in the same direction as the rat when fields were calculated, e.g., moving “downward” with posterior

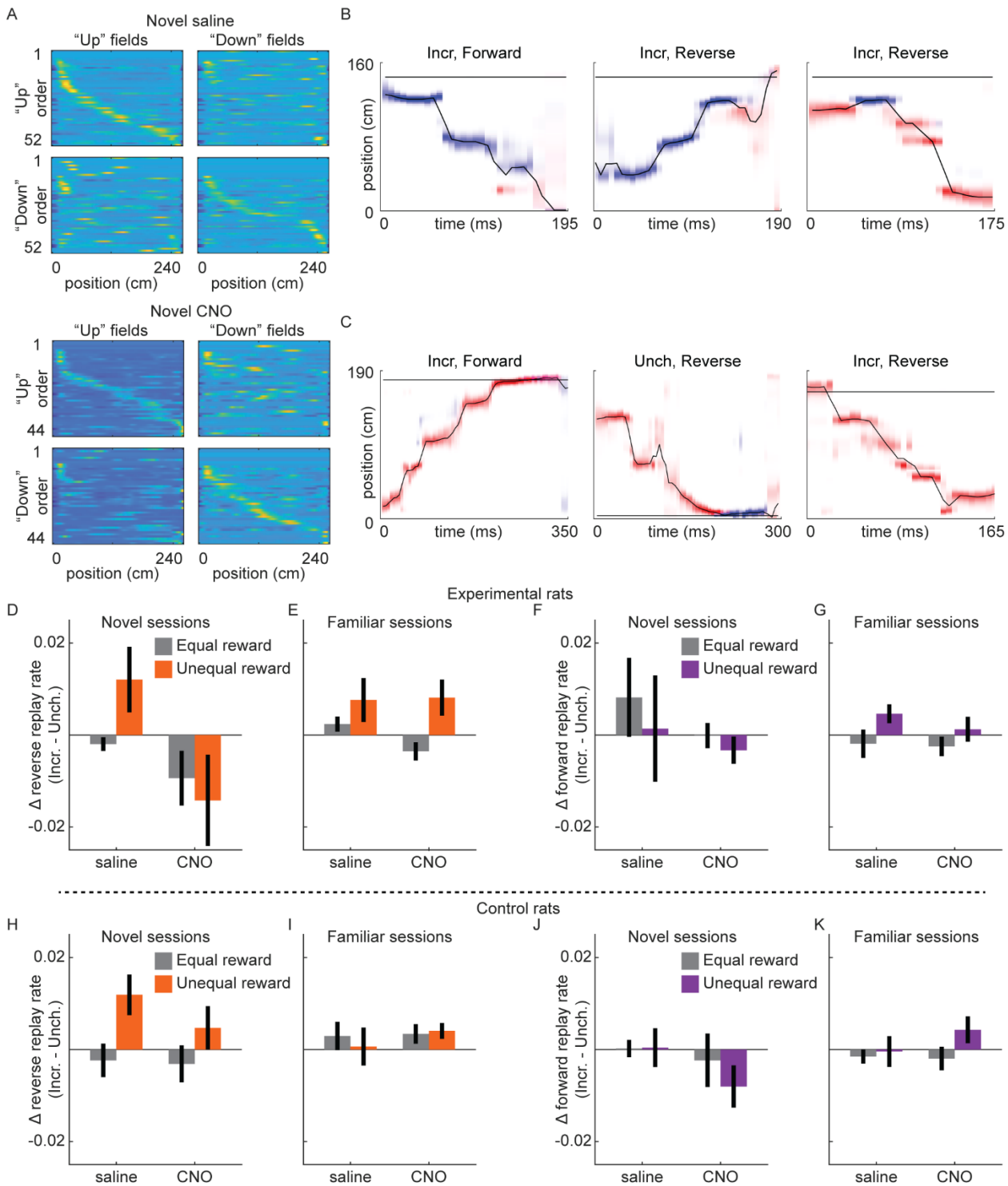


Figure 4. Replay recruitment by reward change in novel sessions requires VTA signaling.

(A) Place cells exhibit directional place fields on the linear track. Fields calculated from movement in a particular direction ("right" fields and "left" fields), ordered based on field center location in either running direction ("right" order and "left" order). Example saline session and CNO session from experimental rat 3. See also Figure S7 and S8. **(B)** Three example replays from Epoch 2 of a novel saline session from experimental rat 3. Red, posterior in upwards map; blue, posterior in downwards map. Title indicates

reward end (Incr., Unch.) and replay direction (Reverse, Forward). The horizontal black line indicates rat position. **(C)** Three example replays from Epoch 2 of a novel CNO session from experimental rat 3, as in (B). **(D)** The difference in rate of reverse replay at each end (Incr. – Unch.) in novel sessions in experimental rats. Error bars, standard error of the mean. Reward condition is indicated by color (equal reward, epoch 1 and 3, gray; unequal reward, epoch 2, orange), and drug condition is indicated on the x-axis. The difference between equal and unequal reward conditions was assessed with a three-way ANOVA with drug, novelty, and replay directionality: drug X novelty X directionality ($F[1,106]=4.64$, $p=0.0335$), all other terms $p>0.05$. **(E)** Same as (D), but for familiar sessions. **(F)** Same as (D), but for forward replay. **(G)** Same as (F), but for familiar sessions. **(H)** Same as (D), but for control rats. The difference between equal and unequal reward conditions was assessed with a three-way ANOVA with drug, novelty, and replay directionality: novelty X directionality ($F[1,101]=9.04$, $p=0.0034$), all other terms $p>0.05$. **(I)** Same as (H), but for familiar sessions. **(J)** Same as (H), but for forward replay. **(K)** Same as (J), but for familiar sessions.

probability in the “downward” place field map (Fig. 4B, left panel). Reverse replays were trajectories that moved in the opposite direction as the rat, e.g., moving “upward” with posterior probability in the “downward” place field map or vice versa (Fig. 4B, middle and right panels). We found forward and reverse replays occurred in all conditions, including in novel sessions with saline (Fig. 4B) or CNO (Fig. 4C). We therefore asked how novelty and drug condition influenced the effect of reward change on rates of reverse and forward replay.

Consistent with previous work (Ambrose et al., 2016), the rate of reverse replay was strongly modulated by the reward volumes on the track. Excluding novel CNO sessions for the moment, in all other conditions in experimental rats, when reward was larger at the Incr. end than the Unch. end (unequal reward), reverse replay was significantly increased at the Incr. end relative to when rewards were equal (novel saline: equal reward, 0.0019±0.0009 replay/s; unequal reward, 0.0121±0.004 replay/s; two-sample t-test, $t(420)=3.235$, $p=0.0013$; familiar saline: equal reward, 0.0033±0.0012 replay/s; unequal reward, 0.0128±0.0025 replay/s; two-sample t-test, $t(907)=3.822$, $p=0.0001$; familiar CNO: equal reward, 0.0033±0.001 replay/s; unequal reward, 0.0139±0.0022 replay/s; two-sample t-test, $t(1153)=5.234$, $p<10^{-6}$). The rate of reverse replay was not significantly increased at the Unch. end with unequal rewards in any of these conditions (all $p>0.05$). This led to a bias for reverse replay to preferentially occur at the Incr. end when rewards were unequal (Fig. 4D-E). In control rats, reward changes caused similar changes to the balance of reverse replay (Fig. 4H-I), with a significantly larger swing in reverse replay bias in novel sessions (three-way ANOVA with drug, novelty, and replay directionality: novelty X directionality, $F[1,101]=9.04$, $p=0.0034$; all other terms, $p>0.05$). Reward changes caused no consistent effects in the rates of forward replay in either animal group (Fig. 4F-G, Fig. 4J-K).

Conversely, in novel CNO sessions in experimental rats, reverse replay rate failed to be biased towards the larger reward location (Fig. 4D). With unequal reward, the rate of reverse replay did increase at the Incr. end (equal reward, 0.0053±0.0016 replay/s; unequal reward, 0.0116±0.0029 replay/s; two-sample t-test, $t(332)=2.043$, $p=0.0419$), but also increased somewhat at the Unch. end (equal reward, 0.0173±0.004 replay/s; unequal reward, 0.0249±0.0068; two-sample t-test, $t(319)=1.019$, $p=0.309$), leading to no consistent change in the difference at the two reward ends. This effect of VTA inactivation on the bias of replay between the reward ends when reward contingencies changed was specific to novel sessions and reverse replay (three-way

ANOVA with drug, novelty, and replay directionality: drug X novelty X directionality, $F[1,106]=4.64$, $p=0.0335$; all other terms, $p>0.05$). VTA dopamine signaling was therefore required to direct reward-related changes in reverse replay, specifically in novel environments.

Discussion

Here we demonstrated a critical role for VTA dopamine signaling in driving hippocampal SWR and reverse replay selectively to locations with increased reward. Surprisingly, we found this was true only in novel environments, with only modest effects of VTA inactivation on SWR rates and no discernible effect on replay rates in environments that had been explored several times before. We additionally recorded activity in a modified task that allowed us to differentiate SWR rate modulation by value and RPE. While SWR rate was modulated by RPE, VTA inactivation had little effect on this RPE-like modulation, suggesting that at least in familiar environments normal VTA dopamine signaling is dispensable for this reward-related hippocampal activity.

Why is VTA inactivation particularly disruptive during novel experiences? Dopamine neuron firing rates are elevated in novel environments (McNamara et al., 2014; Takeuchi et al., 2016). More specifically, early in experience dopamine neuron activity ramps while mice run towards both larger and smaller rewards and this ramping activity declines over experience until modest ramps persist only towards the larger reward (Guru et al., 2020). Activation of VTA projections to dorsal CA1 improves retention of spatial learning of a novel maze configuration, while also promoting replay-related reactivation (McNamara et al., 2014), while inactivation of VTA causes an increase in spatial working memory-related errors in novel, but not familiar, environments (Martig et al., 2009). The results presented here support the hypothesis that VTA is critically involved in learning in new environments, as its inactivation prevents the selective recruitment of replay-associated planning or memory consolidation mechanisms to high value locations.

VTA is not the sole source of dopamine release in hippocampus, with recent work demonstrating locus coeruleus (LC) axons likely provide the bulk of dopamine to dorsal CA1 and can be necessary for novelty-related spatial learning (Kempadoo et al., 2016; Takeuchi et al., 2016). LC axons in CA1 are active in locations immediately preceding a new reward location in a familiar environment but not in a novel one, despite similar behavior in both cases indicating mice had learned the reward locations (Kaufman et al., 2020). This result, coupled with findings that LC neurons are modulated by reward-predicting stimuli similarly to substantia nigra dopamine neurons (Bouret et al., 2012; Bouret and Richmond, 2015), suggests LC activity can convey reward-related information and thereby compensate for VTA inactivation, but only in familiar environments where it is not signaling more general novelty. Altogether, our work adds to the body of evidence that VTA can directly or indirectly mediate hippocampal plasticity and spatial learning and memory (Gasbarri et al., 1996; McNamara et al., 2014; Rosen et al., 2015; Rossato et al., 2009), and suggests an intriguing distinction between the function of VTA and LC dopamine release in hippocampus (Duszkiewicz et al., 2019).

These results also support the hypothesis that reverse replay is intimately involved in reward learning (Ambrose et al., 2016; Foster and Wilson, 2006; Mattar and

Daw, 2018). By activating representations associated with the location of the current reward and then progressing sequentially to earlier positions that preceded reward, reverse replay may provide a neural eligibility trace by which spatial positions can be associated with their proximity to reward (Foster and Wilson, 2006; Sutton and Barto, 1998). Dopamine release at reward detection and consumption would then couple a temporal gradient of dopamine concentration with the temporally-extended, reverse sequential activation of states that led to that reward. Indeed, VTA dopamine neurons are activated when SWR and replay occur during a spatial working memory task, but not during subsequent sleep (Gomperts et al., 2015), indicating close coordination specifically during reward learning. CA1, VTA, and medial prefrontal cortex neurons are jointly coupled via oscillatory mechanisms during spatial working memory (Fujisawa and Buzsáki, 2011), suggesting downstream targets of both replay (Berners-Lee et al., 2021) and VTA dopamine neurons (Lammel et al., 2008) may receive temporally-precise conjunctive input from them. We expect future work aimed at untangling under what conditions VTA and replay influence each other and coordinate to provide downstream areas with sequential activity in the presence of dopamine to be particularly fruitful in understanding how reward drives spatial learning.

Author Contributions

M.R.K. and D.J.F. conceived of and designed the study. M.R.K. acquired the data and performed the analyses. M.R.K. and D.J.F. wrote the manuscript.

Acknowledgements

We thank Stanford Gene Vector and Virus Core and Karl Deisseroth for viral constructs and the Biological Imaging Facility at University of California, Berkeley for assistance with tissue imaging. This work was supported by NIH grant NS113557. Animal use conformed to NIH guidelines and was approved by the UC Berkeley Animal Care and Use Committee.

Declaration of Interests

The authors declare no competing interests.

Methods

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, David Foster (davidfoster@berkeley.edu).

Materials availability

This study did not generate any unique reagents.

Data and code availability

All data and code are available from the Lead Contact upon request. Custom analysis code is publicly available at the DOI listed in the key resource table. Any additional information required to reanalyze the data reported in this work is available from the Lead Contact upon request.

Experimental model and subject details

All experimental procedures were performed in accordance with the University of California Berkeley Animal Care and Use Committee and US National Institutes of Health guidelines. A total of twelve adult male Sprague Dawley TH-cre knock-in rats (inotiv, HsdSage:SD-TH$^{em1(IRES\text{-}Cre)Sage}$, age 3-10 months, 300-750 g) were used in this experiment, of which seven contributed data to the present report. Two were excluded from further analysis due to lack of virus expression evident in post-mortem immunohistochemistry, two were excluded due to faulty recording hardware, and one was excluded due to non-performance of behavioral tasks. Animals were housed on a standard, non-inverted 12-h light cycle. Rats were pair-housed with littermates prior to the start of experiments, after which they were single-housed.

Behavioral pre-training

Adult male Sprague Dawley TH-cre knock-in rats were fed ad lib and handled daily prior to experimental training. They were then food restricted to 85-90% of baseline weight and trained to collect liquid chocolate reward (0.1 ml) from each end of a single linear track (200 cm length) for at least 15 sessions. 3-6 other linear tracks were present in the room during this pre-training, with positions constant for the duration of experiments with each animal.

Surgical procedures

Each rat underwent virus injection and drive implantation in one surgery (control rats 1 and 2) or two surgeries spaced 4-20 days apart (experimental rats 1-4, control rat 3).

Virus injection. All virus was obtained from the Stanford Gene Vector and Virus Core under material transfer agreement with the laboratory of Karl Deisseroth. Experimental rats were injected with AAV-DJ-EF1a-DIO-hM4D(Gi)-mCherry (GVVC-AAV-129) and control rats were injected with AAV-DJ-EF1a-DIO-mCherry (GVVC-AAV-14), with 1 µl of virus delivered stereotactically to VTA in each hemisphere (-5.6 mm posterior, ± 0.7 mm lateral, and -8 mm ventral, all from bregma and skull surface). Data collection began four weeks after virus injection to allow for expression.

Recording drive implantation. Each rat was implanted with a recording microdrive, targeting bilateral (32 tetrodes, ~40 g, n = 4 rats) or unilateral (20 tetrodes, ~35 g, n = 2 rats; 6 tetrodes, ~20 g, n = 1 rat) dorsal CA1. Each tetrode bundle of four platinum iridium wires (Neuralynx) was independently adjustable and electroplated with gold to an impedance of 150-300 kΩ. Tetrodes were advanced over the course of 1-3 weeks to the pyramidal cell layer. Rats were reintroduced to the pre-training linear track after several days of post-surgical recovery.

Tissue processing and immunohistochemistry

Eight weeks after virus injection, rats were deeply anesthetized with isoflurane and transcardially perfused with phosphate-buffered saline (PBS) and then 4% paraformaldehyde (PFA) in PBS. Brains were stored in 4% PFA for >24 hours, then 30% sucrose dissolved in PBS for >7 days for cryoprotection. 20-40 µm sections were made in a cryostat and mounted on slides. For tyrosine hydroxylase (TH) staining, all steps were performed at room temperature in a dark container on a slow orbital shaker. Sections were rinsed three times for 10 minutes each in PBS, then incubated for 2 hours in blocker (3% normal donkey serum and 0.3% Triton-X in PBS). Sections were then kept for 16-20 hours in blocking buffer with primary antibody (1:200, rabbit α-TH, EMD Millipore 657012, or sheep α-TH, Abcam ab113). After three 10-minute washes in PBS, sections were incubated with secondary antibody in blocking buffer for 2 hours (1:200, Alexa Fluor 488-conjugated α-rabbit, Invitrogen ThermoFisher R37118, or Alexa Fluor 488-conjugated α-sheep, Abcam ab150177). Imaging was performed at the Biological Imaging Facility at the University of California Berkeley using a Zeiss AxioImager M2.

Task design

At least 10 minutes prior to beginning a recording session (except for experimental rat 1, with an average of 4 minutes before recording session), rats were injected intraperitoneally (IP) with saline or 1-4 mg/kg clozapine N-oxide (CNO) solution (2-4 mg/ml in diH_2O with 50-100 µl dimethyl sulfoxide). 1-4 sessions were completed each day, with at least 1.5 hours between injections. To prevent the possibility of carry over effects of CNO, saline sessions never followed CNO sessions in the same day (except for 3 recording days in experimental rat 3, when CNO preceded saline sessions by > 4 hours).

In Experiment 1, animals progressed through three epochs. In Epoch 1, animals collected 0.1 ml rewards from each end for 10-20 laps. Then, unsignaled to the rat, the session entered Epoch 2, where reward at one end (Incr. end) was increased to 0.4 ml while the other (Unch. end) remained at 0.1 ml. The assignment of track ends to be Incr. and Unch. randomly varied session to session. After 10-20 laps in Epoch 2, the reward changed again unsignaled to the rat, with both reward ends again delivering 0.1 ml in Epoch 3. Rats completed up to 20 laps in Epoch 3, before being removed and placed back into a rest box. This same task was repeated on distinct linear tracks that varied based on position in the room, material of construction, color, length, orientation, and reward well size and position. Sessions were classified as either “novel” (the 1st or 2nd experience on a particular linear track) or “familiar” (3rd or later experience on a specific track). The track used for pre-training was used first for both saline and CNO sessions.

Then, each novel track was used for 2-6 sessions, with all sessions consisting of only saline or CNO (excluding one track each in experimental rats 2 and 3 that had both saline and CNO sessions). The assignment of saline or CNO to each novel track was varied across rats.

In Experiment 2, reward at the stable end (0.2 ml) remained fixed throughout the session, while at the volatile end it varied pseudorandomly every lap between 0 and 0.8 ml (mean 0.37 ml; blocks of 20 laps were comprised of 3 laps x 0 ml, 4 laps x 0.1 ml, 3 laps x 0.2 ml, 4 laps x 0.4 ml, and 6 laps x 0.8 ml). Rats were allowed to continue running until sated. Which track end was assigned to be stable and volatile varied randomly session by session. Each rat performed saline and CNO sessions of this task on the same linear track. The linear track was initially novel (except for in experimental rat 3). However, 1-2 saline sessions preceded the first CNO session, rendering it familiar for almost all CNO sessions and most saline sessions. In each rat, all experiment 2 sessions were completed after all experiment 1 sessions.

Data acquisition

Rat position was monitored at 30 frames/s using overhead camera and LEDs on the recording drive, then tracked using automated software (Spike Gadgets). Two-dimensional position and velocity were smoothed using a 7-bin median average, followed by a 5 bin Gaussian filter. Linearized position was then used for further analysis. Neural data was collected using a 128-channel wireless HH128 headstage (Spike Gadgets). LFP was sampled at 30 kHz and spikes extracted based on threshold crossing of 40-60 µv (Trodes software, Spike Gadgets). Individual units were differentiated based on manual clustering of spike waveform peak amplitudes using custom software (xclust2, M. A. Wilson, MIT).

Behavioral analysis

Reward end visits were defined as periods when the rat was within 10 cm of the end of the track (approximately ~3-5 cm from the reward well, depending on the track). When analyzing visit durations, we excluded a small number of outliers (< 15 total across all sessions) that were longer than 60 s.

LFP analysis

For each sesson, 2-5 tetrodes with visible sharp wave ripples (SWR) were selected for SWR analysis. LFP from one channel from each tetrode was band-pass filtered between 150 and 250 Hz, then the smoothed (Gaussian kernel, 12 ms s.d.), absolute value of the Hilbert transform was averaged across tetrodes. For detecting SWR, we examined periods when the rat position was within 10 cm of the reward wells with velocity ≤ 8 cm/s. SWR were classified as local peaks when the average ripple power exceeded 4 s.d. above the mean, with start and end points defined as the time ripple power reached the mean before and after the peak, with a minimum start to end duration of 150 ms and maximum of 1 s. SWR rate for each reward end visit was then calculated as the number of SWR detected divided by the total duration of rat velocity ≤ 8 cm/s during that end visit. During Experiment 1, we considered SWR rate during the first 10 s of each end visit to isolate the reward consumption-related period and exclude

occasional longer task-disengaged resting periods. In Experiment 2, we included the first 20 s of each end visit to allow for the longer consumption time required for 0.8 ml.

In Experiment 2, we defined mean-subtracted SWR rate:

$$\text{Mean-subtracted rate } (x,y) = \text{rate } (x,y) - \text{rate } (x)$$

Where x and y are current and previous volatile volumes, respectively.

Single unit analysis

Place fields were calculated for each neuron based on spiking activity while the rat velocity exceeded 8 cm/s. Position was binned into 2 cm bins and directional place fields were calculated as the histogram of spike counts in each position bin (smoothed with Gaussian kernel, 4 cm s.d.) normalized by the animal's occupancy in each bin, separately for periods when the rat was moving in each direction on the linear track (e.g., left and right). We calculated several properties of place fields to determine whether novelty or VTA inactivation affected them. The map direction correlation was defined as the Pearson correlation between the place field calculated in each running direction, such that a value of 1 indicates perfectly reliable firing dependent only on position, not running direction. The lap to session correlation was defined for each neuron only for the running direction with higher max firing rate. For that running direction, a place field was calculated independently for each lap, smoothed (Gaussian kernel, 4 cm s.d.), correlated to the directional field calculated from the entire session, and the average correlation coefficient taken across all laps.

Replay analysis

Candidate replay events were determined based on population activity while the rat was not running (velocity ≤ 8 cm/s) and near the reward wells (10 cm away at most). Population spike density was binned into 1 ms bins and smoothed (Gaussian kernel, 20 ms s.d.) and candidate events defined as local peaks when the population rate exceeded the mean by 3 s.d. and that lasted at least 50 ms, with start and end defined as the nearest times the rate crossed the mean before and after the peak. A memory-less Bayesian decoding algorithm was used to classify both position and running direction during candidate events, as in previous work 13. Replay position in each running direction was estimated in time windows of 40 ms, beginning at the start of the event, and advancing in 5 ms steps. The start/end of putative trajectories within a candidate event were determined by removing bins at either end of the event that contained zero spikes or had a position difference > 50% of the track length from the next/previous window. Candidate events with a remaining length of at least 5 time bins, an absolute weighted correlation (Wu and Foster, 2014) exceeding 0.5, and at least 55% of the posterior probability in one of the running directions (Ambrose et al., 2016) were classified as replay. Replays were classified as forward or reverse by comparing the direction of replay movement across the track with the running direction map containing the majority of the posterior probability: if they matched (e.g., the replay moved upward and used upward fields), the replay was classified as forward, and otherwise it was classified as reverse.

Only sessions with sufficiently accurate behavioral position decoding accuracy during run were included. Bayesian decoding using the directional place fields was applied to 250 ms non-overlapping windows covering the entirety of each session. For

all time bins with mean animal velocity >20 cm/s, position >20 cm from reward wells, and >5 total spikes from any neurons, actual and decoded position and running direction were compared, yielding a position decoding error (distance in cm) and direction decoding match (same or different). Sessions with mean decoding error >35 cm or direction match <60% were excluded.

Statistics

A mixed effects Poisson generalized linear model was used to test which experimental variables affected SWR rate using the Matlab *fitglme* function (Mathworks), similarly to previous work 13. Animal ID was modeled as a random effect, allowing baseline SWR rate to vary across rats. The full model was as follows:

SWR rate = exp[b0 + b1 × (Incr. reward end) + b2 × (Epoch 2) + b3 × (CNO) + b4 × (Incr. reward end) × (Epoch 2) + b5 ×(Incr. reward end) × (CNO) + b6 × (Epoch 2) × (CNO) + b7 × (Incr. reward end) × (Epoch 2) × (CNO)]

Where “Incr. reward end” is a dummy variable indicating the rat is at the Incr. reward end, “Epoch 2” is a dummy variable indicating the visit is occurring in Epoch 2, and “CNO” is a dummy variable indicating it is a session with CNO injected. The coefficient for each term corresponds to the log multiplicative change in SWR rate from the reference condition (animal-specific rate at Unch. end, not in Epoch 2, of a saline session). The offset term was log(duration) of each stopping period, so the model fit SWR rate, rather than SWR count. Experimental and control rats were fit separately with this model, as were familiar and novel sessions.

Bootstrapping was used to assess the effect of drug on SWR rate in each experimental condition. For each combination of novelty, reward end, and epoch, drug identity was shuffled 5,000 times, generating a distribution of the chance difference between the SWR rate in saline vs. CNO sessions. P-values were determined using one-tailed tests, under the hypothesis that CNO would cause lower SWR rates at the Incr. end and higher SWR rates at the Unch. end when compared to saline.

In ANOVA used to assess the effect of various experimental variables on behavioral and neural measurements, the following variables were consistently defined: “animal group” was a dummy variable indicating experimental rats, “drug” was a dummy variable indicating CNO, “novelty” was a dummy variable indicating novel session, “epoch” was a categorical variable indicating epoch number, “reward end” was a dummy variable indicating Incr. reward end, “RPE sign” was a dummy variable indicating a positive RPE, and “previous volume” was a categorical variable indicating the volatile volume at the previous visit.

Supplemental Figure Legends

Figure S1. Behavioral effects of novelty and VTA inactivation. Related to Figure 1. **(A)** In novel sessions, Unch. visit duration decreased from Epoch 1 to Epoch 2, while CNO additionally led to longer visit duration in experimental rats. Mean ± standard error, Exp Saline, Epoch 1: 7.181±0.33, Epoch 2: 4.892±0.2, two sample t-test: t(348)=5.836, $p<10^{-5}$; Exp CNO, Epoch 1: 10.542±0.65, Epoch 2: 6.789±0.41, two sample t-test: t(297)=5.012, $p<10^{-5}$. Con Saline, Epoch 1: 7.594±0.28, Epoch 2: 4.8±0.14, two sample t-test: t(401)=9.054, $p<10^{-10}$; Con CNO, Epoch 1: 7.267±0.26, Epoch 2: 4.834±0.12, two sample t-test: t(426)=8.471, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch (F[1,1472]=171.66, $p<10^{-10}$), drug (F[1,1472]=33.3, $p<10^{-5}$), group (F[1,1472]=32.57, $p<10^{-5}$), drug X group (F[1,1472]=41.66, $p<10^{-5}$), epoch X drug X group (F[1,1472]=4.5, p=0.034). **(B)** In novel sessions, Incr. visit duration increased from Epoch 1 to Epoch 2, while CNO additionally led to longer visit duration in experimental rats. Mean ± standard error, Exp Saline, Epoch 1: 7.179±0.39, Epoch 2: 9.968±0.3, two sample t-test: t(343)=-5.668, $p<10^{-6}$; Exp CNO, Epoch 1: 10.16±0.74, Epoch 2: 13.721±0.48, two sample t-test: t(293)=-4.164, p=0.00004. Con Saline, Epoch 1: 7.478±0.29, Epoch 2: 10.907±0.24, two sample t-test: t(395)=-9.086, $p<10^{-10}$; Con CNO, Epoch 1: 6.65±0.21, Epoch 2: 10.506±0.15, two sample t-test: t(420)=-15.18, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch (F[1,1451]=192.37, $p<10^{-10}$), drug (F[1,1451]=31.38, $p<10^{-5}$), group (F[1,1451]=31.16, $p<10^{-5}$), drug X group (F[1,1451]=65.62, $p<10^{-10}$). **(C)** In familiar sessions, Unch. visit duration decreased from Epoch 1 to Epoch 2, with only a modest effect of CNO compared to novel sessions. Mean ± standard error, Exp Saline, Epoch 1: 6.011±0.2, Epoch 2: 4.7±0.18, two sample t-test: t(1180)=4.806, $p<10^{-4}$; Exp CNO, Epoch 1: 6.037±0.18, Epoch 2: 5.371±0.17, two sample t-test: t(1053)=2.712, p=0.0068. Con Saline, Epoch 1: 5.969±0.17, Epoch 2: 4.465±0.13, two sample t-test: t(746)=7.057, $p<10^{-10}$; Con CNO, Epoch 1: 6.035±0.13, Epoch 2: 4.174±0.07, two sample t-test: t(858)=13.165, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch (F[1,3837]=120.87, $p<10^{-10}$), group (F[1,3837]=9.22, p=0.0024), epoch X group (F[1,3837]=8.15, p=0.0043), epoch X drug X group (F[1,3837]=4.26, p=0.0391). **(D)** In familiar sessions, Incr. visit duration increased from Epoch 1 to Epoch 2. Mean ± standard error, Exp Saline, Epoch 1: 6.058±0.19, Epoch 2: 10.463±0.24, two sample t-test: t(1169)=-14.293, $p<10^{-10}$; Exp CNO, Epoch 1: 5.77±0.18, Epoch 2: 11.07±0.29, two sample t-test: t(1045)=-15.654, $p<10^{-10}$. Con Saline, Epoch 1: 6.519±0.2, Epoch 2: 11.12±0.23, two sample t-test: t(741)=-14.979, $p<10^{-10}$; Con CNO, Epoch 1: 6.018±0.15, Epoch 2: 10.206±0.15, two sample t-test: t(852)=-19.849, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch (F[1,3807]=885.39, $p<10^{-10}$), drug X group (F[1,3807]=7.78, p=0.0053), epoch X drug X group (F[1,3807]=4.44, p=0.0352). **(E)** Unch. visit duration increased from Epoch 2 to Epoch 3. Mean ± standard error, Exp Saline, Epoch 2: 4.743±0.15, Epoch 3: 7.274±0.23, two sample t-test: t(1354)=-9.542, $p<10^{-10}$; Exp CNO, Epoch 2: 5.705±0.16, Epoch 3: 6.898±0.27, two sample t-test: t(1096)=-4.031, $p=10^{-5}$. Con Saline, Epoch 2: 4.58±0.1, Epoch 3: 6.05±0.18, two sample t-test: t(939)=-7.838, $p<10^{-10}$; Con CNO, Epoch 2: 4.391±0.06, Epoch 3: 6.049±0.13, two sample t-test: t(1033)=-12.818, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch (F[1,4422]=196.47, $p<10^{-10}$), group (F[1,4422]=52.77, $p<10^{-5}$),

epoch X drug (F[1,4422]=5.53, p=0.0187), epoch X drug X group (F[1,4422]=9.74, p=0.0018). **(F)** Incr. visit duration decreased from Epoch 2 to Epoch 3. Mean ± standard error, Exp Saline, Epoch 2: 10.351±0.2, Epoch 3: 7.878±0.31, two sample t-test: t(1369)=6.993, $p<10^{-10}$; Exp CNO, Epoch 2: 11.691±0.25, Epoch 3: 7.718±0.35, two sample t-test: t(1116)=9.437, $p<10^{-10}$. Con Saline, Epoch 2: 11.047±0.17, Epoch 3: 6.578±0.18, two sample t-test: t(958)=17.166, $p<10^{-10}$; Con CNO, Epoch 2: 10.304±0.11, Epoch 3: 6.296±0.17, two sample t-test: t(1057)=20.98, $p<10^{-10}$. Three-way ANOVA with epoch, drug, and animal group: epoch (F[1,4500]=491.46, $p<10^{-10}$), group (F[1,4500]=25.7, $p<10^{-5}$), epoch X group (F[1,4500]=9.11, p=0.0026), drug X group (F[1,4500]=10.72, p=0.0011), epoch X drug X group (F[1,4500]=8.49, p=0.0036).

Figure S2. Effect of reward change on running velocity. Related to Figure 1.
Running speed towards the Incr. end in Epoch 2 was consistently significantly faster than towards the Unch. end, across all animal group, drug, and novelty conditions (one-sample t-test: exp, novel, CNO: t(9)=3.96, p=0.003; exp, novel, saline: t(8)=3.45, p=0.009; exp, familiar, CNO: t(33)=6.9, $p<10^{-7}$; exp, familiar, saline: t(35)=3.96, $p<10^{-3}$; control, novel, CNO: t(11)=3.34, p=0.007; control, novel, saline: t(11)=4.26, p=0.001; control, familiar, CNO: t(24)=8.46, $p<10^{-7}$; control, familiar, saline: t(22)=5.6, $p<10^{-4}$). Filled symbol, saline; unfilled symbol, CNO. Bars are standard error.

Figure S3. Modulation of SWR rate by reward increase. Related to Figure 2.
(A) In experimental rats, a mixed effects Poisson GLM was fit to the data and 5,000 drug identity shuffles. The difference between model-predicted SWR rate in saline and CNO sessions at each reward end (Unch. top row, Incr. bottom row) and novelty condition (familiar left column, novel right column), in data (red lines) and in bootstrap shuffles (histogram). Significance values reflect one-tailed hypothesis test, with hypotheses that Unch. saline < Unch. CNO and Incr. saline > Incr. CNO. **(B)** A mixed effects GLM with bootstrap, as in (A), but for control animals.

Figure S4. SWR rate in Epoch 3. Related to Figure 2.
(A) SWR rate as a function of time in stopping period in Epoch 2 and 3 for four example sessions in experimental rats, as in Figure 2a. Epoch 2 (red lines), Epoch 3 (dashed gray lines). SWR rate was binned in 0.25 s windows and smoothed with a 2 bin Gaussian. Line, mean; shading, standard error. **(B)** Same as Figure 2F, but for Epoch 3. **(C)** Same as Figure 2G, but for Epoch 3.

Figure S5. SWR rate at stable end in experimental rats. Related to Figure 3.
(A) At stable end visits in saline sessions, SWR rate was not significantly modulated by the previous volatile end visit reward volume. Pearson correlation between SWR rate and previous volatile volume, r=-0.0643, p=0.21. Two sample t-test between volatile volume ≤ 2 and volatile volume > 2, t(380)=1.465, p=0.144. **(B)** At stable end visits in CNO sessions, SWR rate was not significantly modulated by the previous volatile end visit reward volume. Pearson correlation between SWR rate and previous volatile volume, r=-0.0645, p=0.205. Two sample t-test between volatile volume ≤ 2 and volatile volume > 2, t(386)=1.137, p = 0.256. Two-way ANOVA with drug and previous volatile

volume ≤ 2: drug (F[1,766]=6.43, p=0.0114), volume ≤ 2 (F[1,766]=3.36, p=0.067), drug X volume (F[1,766]=0.03, p=0.853). Error bars, standard error.

Figure S6. SWR rate in all sessions in volatile reward task. Related to Figure 3.
(A) SWR rate as a function of reward volume and time in end visit, as in Figure 3B, for all sessions combined (including saline and CNO sessions in experimental and control rats). Left panel, stable reward end. Right panel, volatile reward end. In stable panel, traces are colored based on previous volatile end visit volume. In volatile panel, traces are colored based on current volatile volume. **(B)** SWR rate at volatile end as a function of current and previous volatile volume, as in Figure 3D, for all volatile reward task sessions. **(C)** SWR rate for each non-zero volatile volume plotted as a function of previous volume, with the mean SWR rate for that current volume subtracted. Unfilled symbols, mean of previous volume across all current volumes. Thick dashed line, linear fit to mean values. Pearson correlation between (ripple rate – mean) and previous volume, r=-0.07, p=0.0014, consistent with RPE coding. Error bars, standard error. **(D)** Positive RPE caused significantly greater ripple rate than negative RPE (two-sample t-test, t[1661]=2.741, p=0.0062). **(E)** SWR rate at the stable end was significantly negatively correlated with the most recent volatile volume (r=-0.06, p=0.003). **(F)** SWR rate at the stable end was significantly greater when the most recent volatile end volume was less than or equal in volume (≤ 2) than when it was greater (two-sample t-test, t[2485]=2.582, p=0.01). **(G)** SWR rate at the volatile end was significantly higher if recent reward history was lower than the average. Reward volume at the 3 previous visits was averaged, then split above and below the median. Poisson GLM with two terms, current volume and reward history (above/below median): current volume, z=22.21, $p<10^{-10}$; history, z=-2.03, p=0.042).

Figure S7. Effect of novelty and VTA inactivation on place cell properties. Related to Figure 4.
(A) Correlation between single lap place fields and session averaged field. Three-way ANOVA with drug, novelty, and animal group: novelty (F[1,3249]=6.75, p=0.0094), novelty X group (F[1,3249]=15.76, p=0.0001), all others, p>0.2. **(B)** Correlation between unidirectional fields calculated separately in each running direction. Three-way ANOVA with drug, novelty, and animal group: drug (F[1,2816]=5.76, p=0.0164), novelty (F[1,2816]=28.21, $p<10^{-10}$), drug X novelty (F[1,2816]=5.52, p=0.0188), novelty X group (F[1,2816]=6.56, p=0.011), all others, p>0.17.

Figure S8. Run decoding accuracy in replay analysis sessions. Related to Figure 4.
(A) Mean decoding error during run. Position and running direction were decoded during periods of strong locomotion (animal velocity >20 cm/s and position >20 cm from the reward wells) in 250 ms bins. Sessions with >35 cm mean decoding error were excluded from analysis. Filled and unfilled symbols are saline and CNO sessions, respectively. Error bars, standard error. Three-way ANOVA with animal group, drug, and novelty, all terms n.s. **(B)** Mean fraction of bins where actual and decoded running

direction were the same. Sessions with <60% match were excluded from analysis. Symbols as in (A). Three-way ANOVA with animal group, drug, and novelty, all terms n.s.

Figure S1

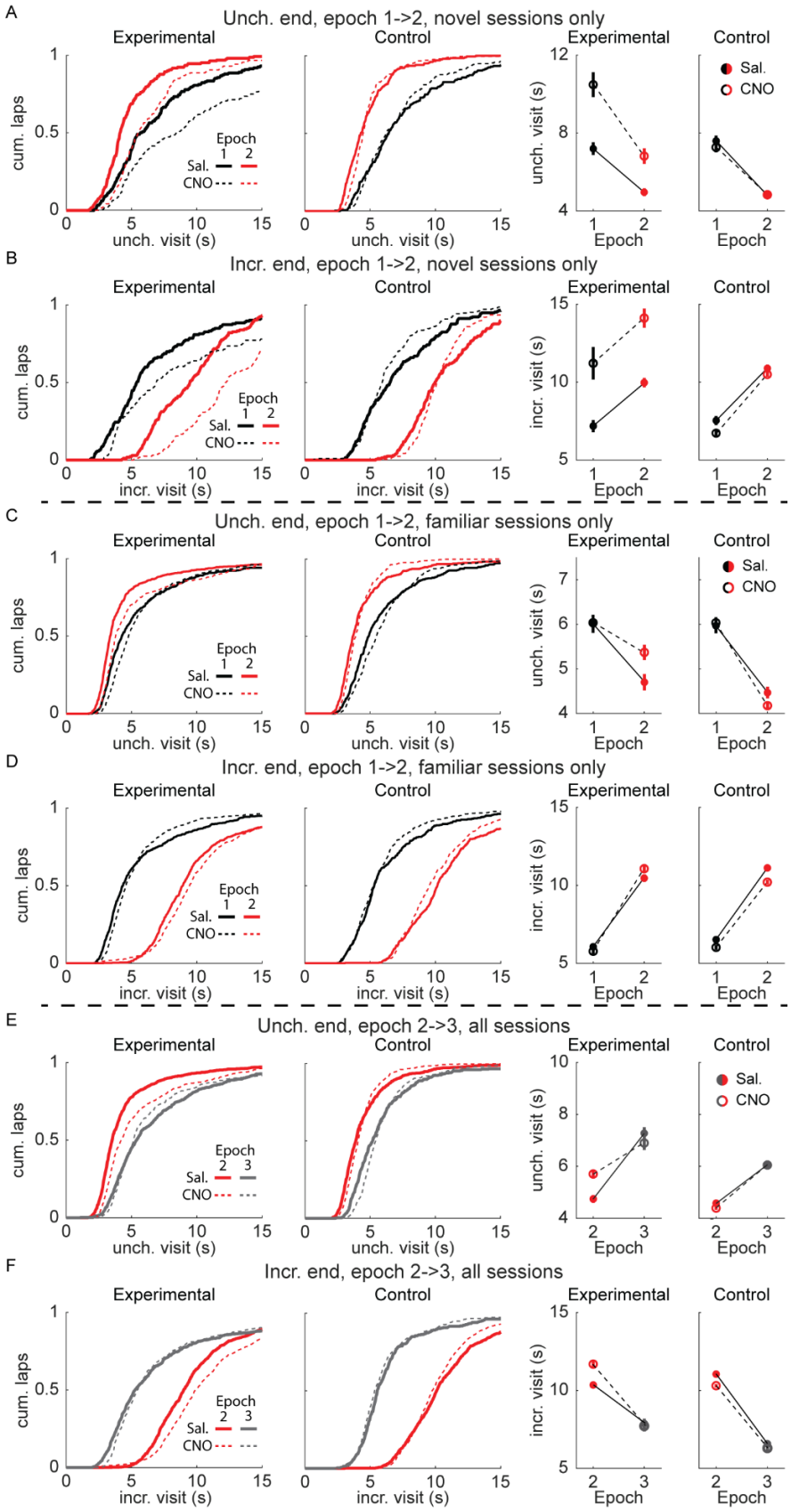

A
Unch. end, epoch 1->2, novel sessions only
Experimental
Control
Experimental
Control
cum. laps
unch. visit (s)
Epoch
1 2
Sal.
CNO
Sal.
CNO
B
Incr. end, epoch 1->2, novel sessions only
incr. visit (s)
C
Unch. end, epoch 1->2, familiar sessions only
D
Incr. end, epoch 1->2, familiar sessions only
E
Unch. end, epoch 2->3, all sessions
2 3
F
Incr. end, epoch 2->3, all sessions

Figure S2

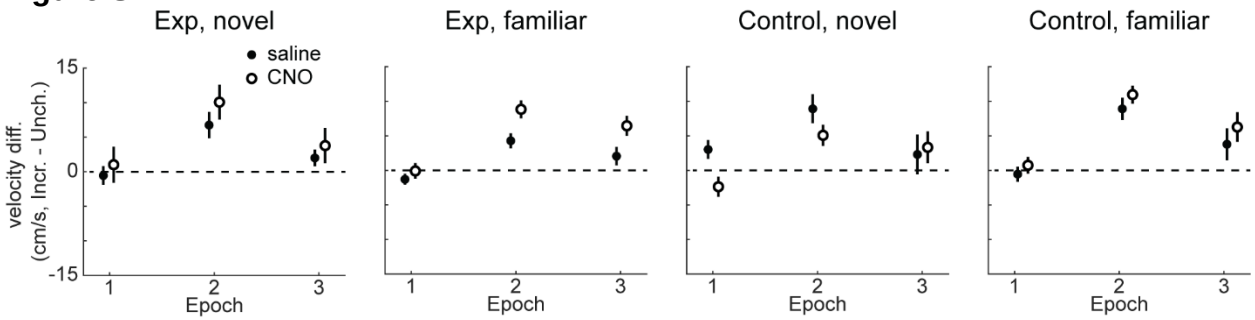

Exp, novel
Exp, familiar
Control, novel
Control, familiar
saline
CNO
velocity diff.
(cm/s, Incr. - Unch.)
15
0
-15
1
2
3
Epoch

Figure S3

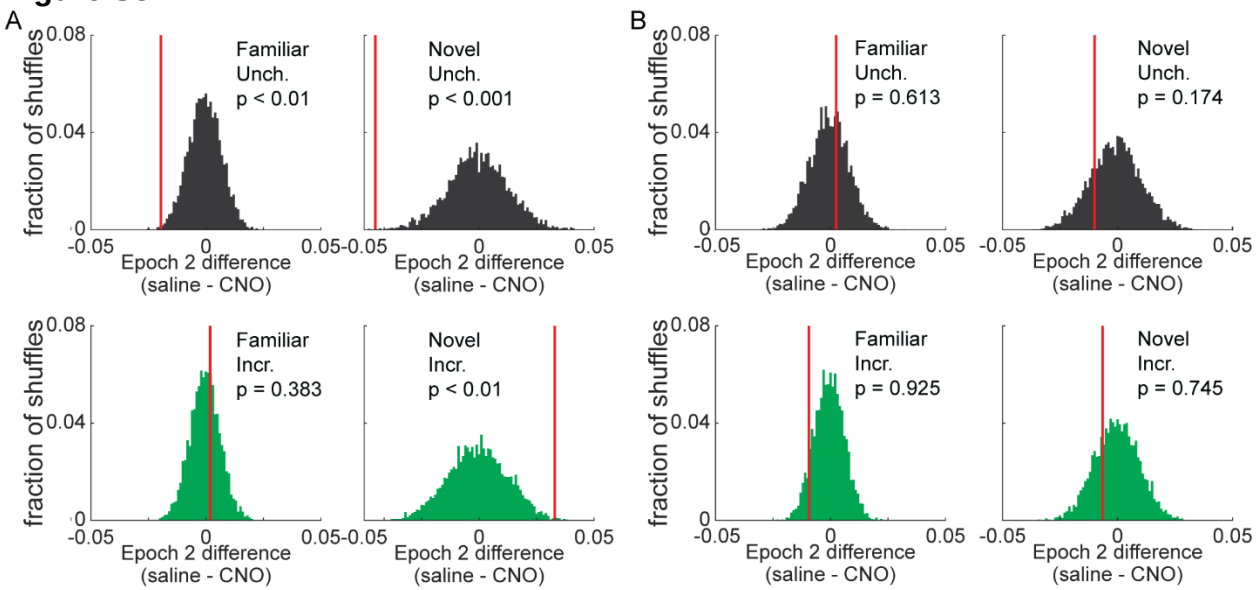

A
B
fraction of shuffles
0.08
0.04
0
-0.05
0
0.05
Epoch 2 difference
(saline - CNO)
Familiar
Unch.
p < 0.01
Novel
Unch.
p < 0.001
Familiar
Unch.
p = 0.613
Novel
Unch.
p = 0.174
Familiar
Incr.
p = 0.383
Novel
Incr.
p < 0.01
Familiar
Incr.
p = 0.925
Novel
Incr.
p = 0.745

Figure S4

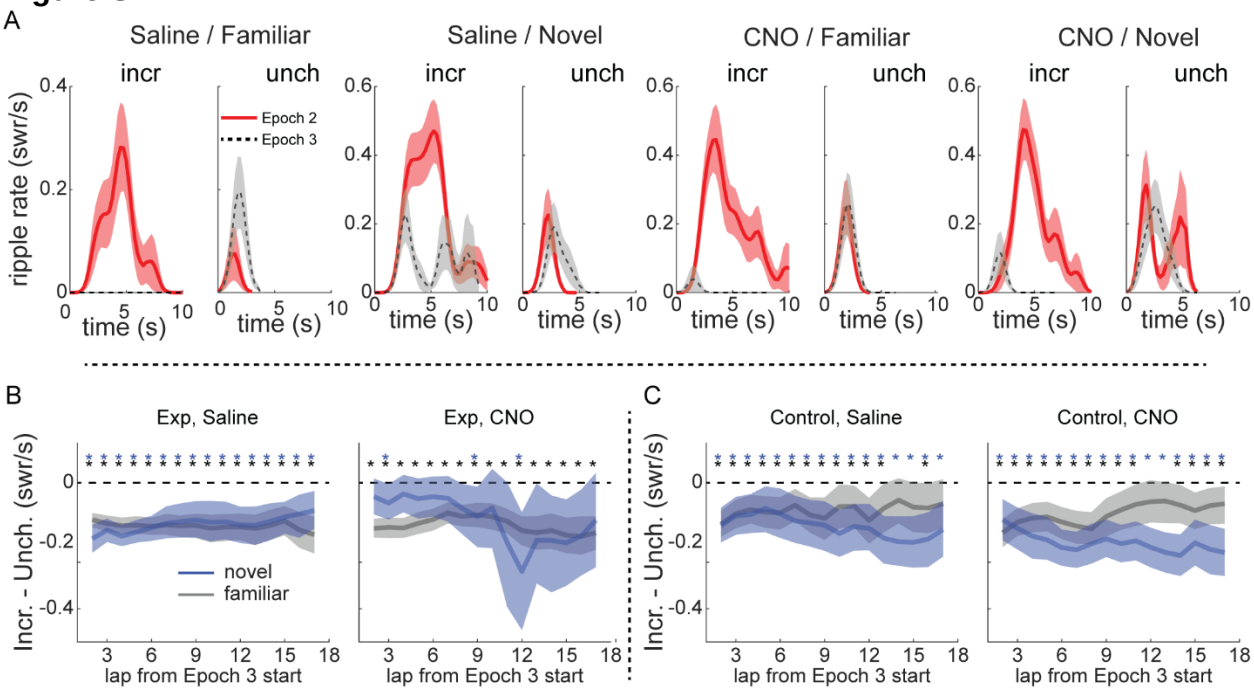

A
Saline / Familiar
Saline / Novel
CNO / Familiar
CNO / Novel
incr
unch
ripple rate (swr/s)
Epoch 2
Epoch 3
time (s)
B
C
Exp, Saline
Exp, CNO
Control, Saline
Control, CNO
Incr. - Unch. (swr/s)
novel
familiar
lap from Epoch 3 start

Figure S5

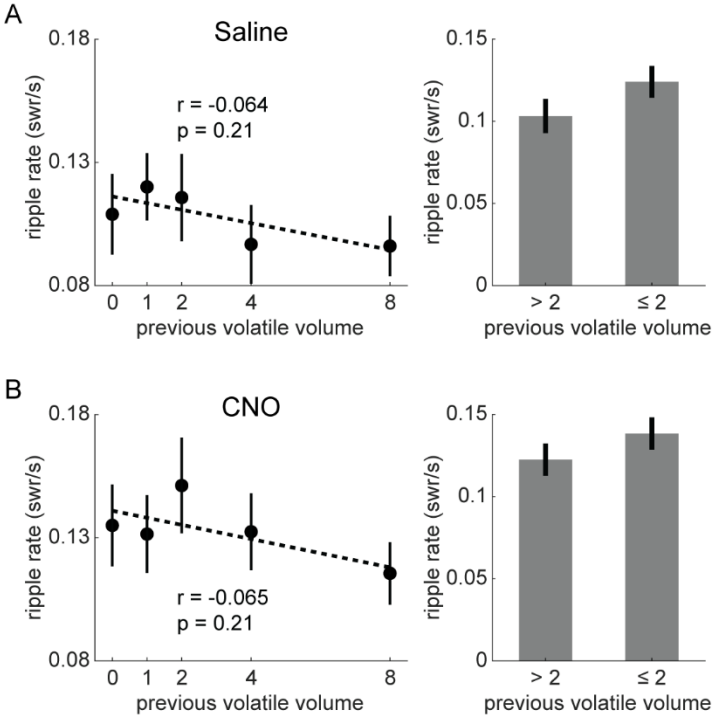

A
Saline
0.18
0.13
0.08
ripple rate (swr/s)
r = -0.064
p = 0.21
0 1 2 4 8
previous volatile volume
0.15
0.1
0.05
0
ripple rate (swr/s)
> 2
≤ 2
previous volatile volume
B
CNO
0.18
0.13
0.08
ripple rate (swr/s)
r = -0.065
p = 0.21
0 1 2 4 8
previous volatile volume
0.15
0.1
0.05
0
ripple rate (swr/s)
> 2
≤ 2
previous volatile volume

Figure S6

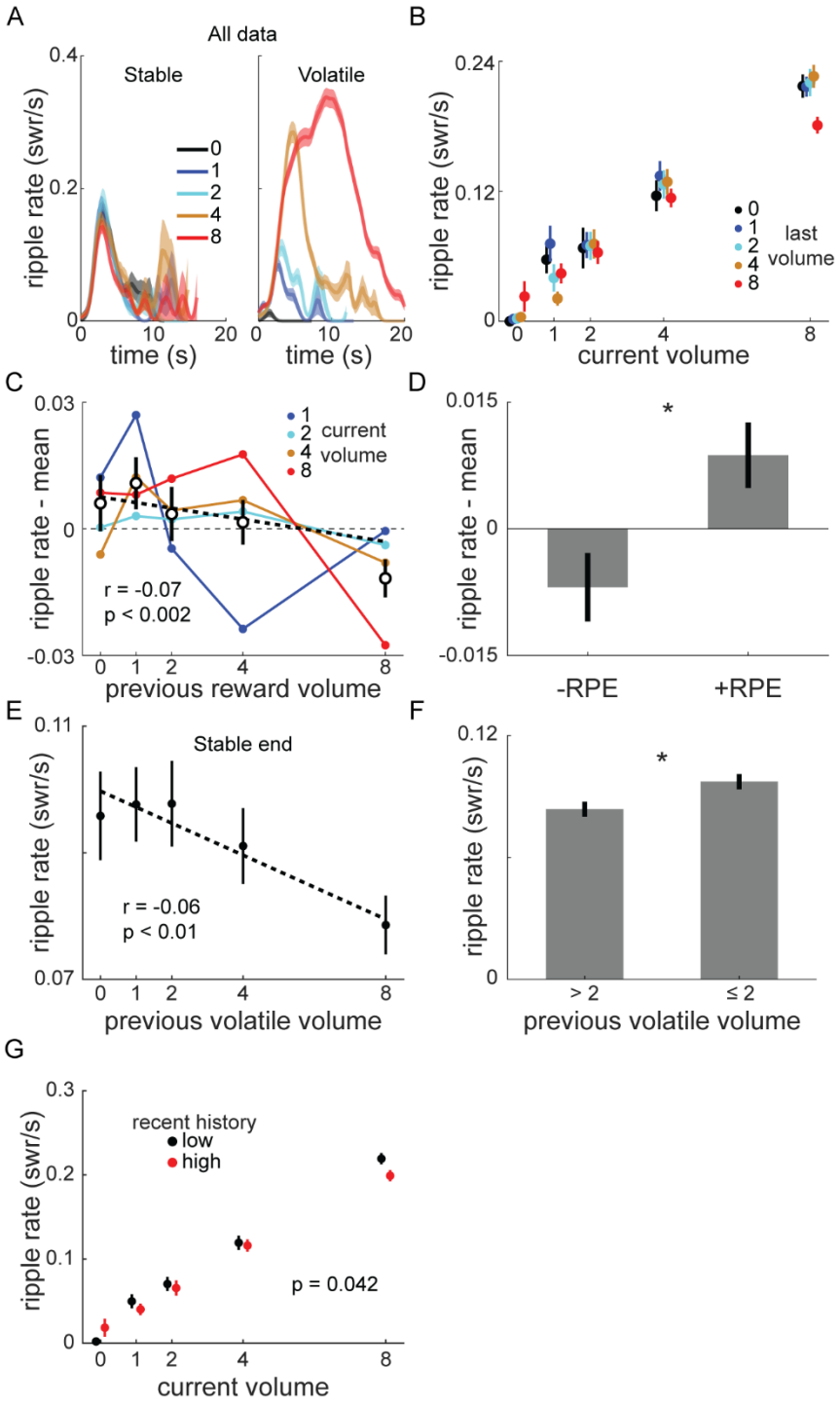

A
All data
Stable
Volatile
ripple rate (swr/s)
0.4
0.2
0
0
10
20
time (s)
0
1
2
4
8
B
ripple rate (swr/s)
0.24
0.12
0
0
1
2
4
8
current volume
last volume
C
ripple rate - mean
0.03
0
-0.03
r = -0.07
p < 0.002
previous reward volume
current volume
D
ripple rate - mean
0.015
0
-0.015
*
-RPE
+RPE
E
Stable end
ripple rate (swr/s)
0.11
0.07
r = -0.06
p < 0.01
previous volatile volume
F
ripple rate (swr/s)
0.12
0
*
> 2
≤ 2
previous volatile volume
G
ripple rate (swr/s)
0.3
0.2
0.1
0
recent history
low
high
p = 0.042
current volume

Figure S7

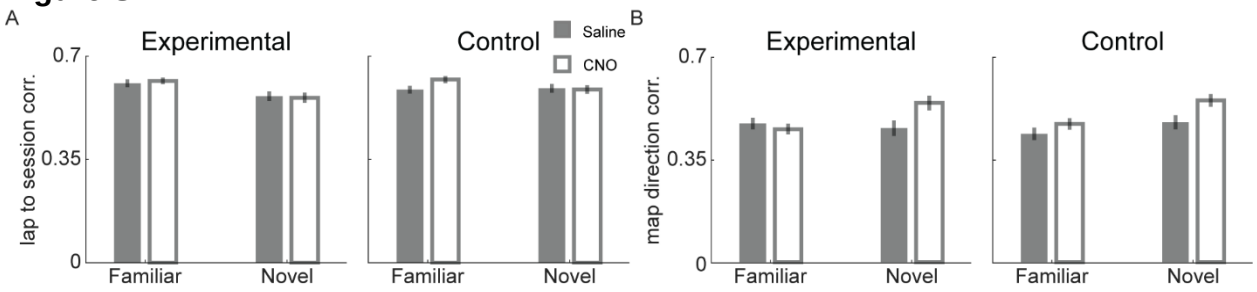
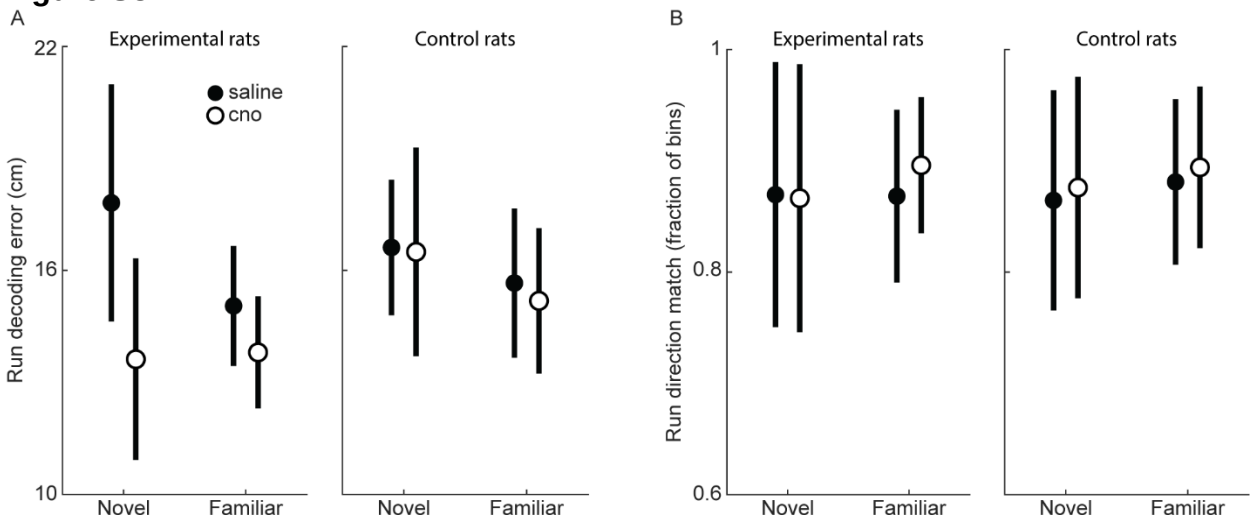

A
Experimental
Control
Saline
CNO
B
Experimental
Control
lap to session corr.
map direction corr.
0.7
0.35
0
Familiar
Novel

Figure S8


A
Experimental rats
Control rats
saline
cno
Run decoding error (cm)
22
16
10
Novel
Familiar
Novel
Familiar
B
Experimental rats
Control rats
Run direction match (fraction of bins)
1
0.8
0.6
Novel
Familiar
Novel
Familiar

References

Ambrose RE, Pfeiffer BE, Foster DJ. 2016. Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward. *Neuron* **91**:1124–1136. doi:10.1016/j.neuron.2016.07.047

Armbruster BN, Li X, Pausch MH, Herlitze S, Roth BL. 2007. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proceedings of the National Academy of Sciences* **104**:5163–5168. doi:10.1073/pnas.0700293104

Berners-Lee A, Feng T, Silva D, Wu X, Ambrose ER, Pfeiffer BE, Foster DJ. 2022. Hippocampal replays appear after a single experience and incorporate greater detail with more experience. *Neuron* **110**:1829-1842.e5. doi:10.1016/j.neuron.2022.03.010

Berners-Lee A, Wu X, Foster DJ. 2021. Prefrontal Cortical Neurons Are Selective for Non-Local Hippocampal Representations during Replay and Behavior. *The Journal of Neuroscience* **41**:5894–5908. doi:10.1523/JNEUROSCI.1158-20.2021

Bouret S, Ravel S, Richmond BJ. 2012. Complementary neural correlates of motivation in dopaminergic and noradrenergic neurons of monkeys. *Front Behav Neurosci* **6**:40. doi:10.3389/fnbeh.2012.00040

Bouret S, Richmond BJ. 2015. Sensitivity of Locus Ceruleus Neurons to Reward Value for Goal-Directed Actions. *The Journal of Neuroscience* **35**:4005–4014. doi:10.1523/JNEUROSCI.4553-14.2015

Buzsáki G. 2015. Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus* **25**:1073–1188. doi:10.1002/hipo.22488

Diba K, Buzsáki G. 2007. Forward and reverse hippocampal place-cell sequences during ripples. *Nat Neurosci* **10**:1241–2. doi:10.1038/nn1961

Duszkiewicz AJ, McNamara CG, Takeuchi T, Genzel L. 2019. Novelty and Dopaminergic Modulation of Memory Persistence: A Tale of Two Systems. *Trends Neurosci*. doi:10.1016/j.tins.2018.10.002

Fernández-Ruiz A, Oliva A, Fermino de Oliveira E, Rocha-Almeida F, Tingley D, Buzsáki G. 2019. Long-duration hippocampal sharp wave ripples improve memory. *Science (1979)* **364**:1082–1086. doi:10.1126/science.aax0758

Fields HL, Hjelmstad GO, Margolis EB, Nicola SM. 2007. Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. *Annu Rev Neurosci* **30**:289–316. doi:10.1146/annurev.neuro.30.051606.094341

Foster DJ. 2017. Replay Comes of Age. *Annu Rev Neurosci* **40**:581–602. doi:10.1146/annurev-neuro-072116-031538

Foster DJ, Wilson M a. 2006. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440**:680–3. doi:10.1038/nature04587

Fujisawa S, Buzsáki G. 2011. A 4 Hz oscillation adaptively synchronizes prefrontal, VTA, and hippocampal activities. *Neuron* **72**:153–65. doi:10.1016/j.neuron.2011.08.018

Gasbarri A, Sulli A, Innocenzi R, Pacitti C, Brioni JD. 1996. Spatial memory impairment induced by lesion of the mesohippocampal dopaminergic system in the rat. *Neuroscience* **74**:1037–1044. doi:10.1016/S0306-4522(96)00202-3

Gomperts SN, Kloosterman F, Wilson M a. 2015. VTA neurons coordinate with the hippocampal reactivation of spatial experience. *Elife* **4**:1–22. doi:10.7554/eLife.05360

Guru A, Seo C, Post RJ, Kullakanda DS, Schaffer JA, Warden MR. 2020. Ramping activity in midbrain dopamine neurons signifies the use of a cognitive map. *BioRxiv*. doi:doi.org/10.1101/2020.05.21.108886

Howe MW, Tierney PL, Sandberg SG, Phillips PEM, Graybiel AM. 2013. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* **500**:575–579. doi:10.1038/nature12475

Jadhav SP, Kemere C, German PW, Frank LM. 2012. Awake hippocampal sharp-wave ripples support spatial memory. *Science (1979)* **336**:1454–8. doi:10.1126/science.1217230

Kaufman AM, Geiller T, Losonczy A. 2020. A Role for the Locus Coeruleus in Hippocampal CA1 Place Cell Reorganization during Spatial Reward Learning. *Neuron* **105**:1018-1026.e4. doi:10.1016/j.neuron.2019.12.029

Kempadoo KA, Mosharov E V, Choi SJ, Sulzer D, Kandel ER. 2016. Dopamine release from the locus coeruleus to the dorsal hippocampus promotes spatial learning and memory. *Proc Natl Acad Sci U S A* **113**:201616515. doi:10.1073/pnas.1616515114

Kentros CG, Agnihotri NT, Streater S, Hawkins RD, Kandel ER. 2004. Increased attention to spatial context increases both place field stability and spatial memory. *Neuron* **42**:283–295. doi:10.1016/S0896-6273(04)00192-8

Kim HGR, Malik AN, Mikhael JG, Bech P, Tsutsui-Kimura I, Sun F, Zhang Y, Li Y, Watabe-Uchida M, Gershman SJ, Uchida N. 2020. A Unified Framework for Dopamine Signals across Timescales. *Cell* **183**:1600-1616.e25. doi:10.1016/j.cell.2020.11.013

Lammel S, Hetzel A, Häckel O, Jones I, Liss B, Roeper J. 2008. Unique Properties of Mesoprefrontal Neurons within a Dual Mesocorticolimbic Dopamine System. *Neuron* **57**:760–773. doi:10.1016/j.neuron.2008.01.022

Li S, Cullen WK, Anwyl R, Rowan MJ. 2003. Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty. *Nat Neurosci* **6**:526–531. doi:10.1038/nn1049

Lisman JE, Grace A a. 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* **46**:703–13. doi:10.1016/j.neuron.2005.05.002

Martig AK, Jones GL, Smith KE, Mizumori SJY. 2009. Context dependent effects of ventral tegmental area inactivation on spatial working memory. *Behavioural Brain Research* **203**:316–320. doi:10.1016/j.bbr.2009.05.008

Mattar MG, Daw ND. 2018. Prioritized memory access explains planning and hippocampal replay. *Nat Neurosci* **21**:1609–1617. doi:10.1038/s41593-018-0232-z

McNamara CG, Tejero-Cantero Á, Trouche S, Campo-Urriza N, Dupret D. 2014. Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. *Nat Neurosci* **17**:1658–1660. doi:10.1038/nn.3843

O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* **34**:171–175. doi:10.1016/0006-8993(71)90358-1

O'Keefe J, Nadel L. 1978. The hippocampus as a cognitive map. Clarendon Press.

Pfeiffer BE, Foster DJ. 2013. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**:1–8. doi:10.1038/nature12112

Rosen ZB, Cheung S, Siegelbaum SA. 2015. Midbrain dopamine neurons bidirectionally regulate CA3-CA1 synaptic drive. *Nat Neurosci* **18**:1763–1771. doi:10.1038/nn.4152

Rossato JI, Bevilaqua LRM, Izquierdo I, Medina JH, Cammarota M. 2009. Dopamine Controls Persistence of Long-Term Memory Storage. *Science (1979)* **325**:1017–1020. doi:10.1126/science.1172545

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science (1979)* **275**:1593–1599. doi:10.1126/science.275.5306.1593

Singer AC, Frank LM. 2009. Rewarded Outcomes Enhance Reactivation of Experience in the Hippocampus. *Neuron* **64**:910–921. doi:10.1016/j.neuron.2009.11.016

Sutton RS, Barto AG. 1998. Reinforcement learning: an introduction. MIT Press.

Takeuchi T, Duszkiewicz AJ, Sonneborn A, Spooner P a., Yamasaki M, Watanabe M, Smith CC, Fernández G, Deisseroth K, Greene RW, Morris RGM. 2016. Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature* 1–18. doi:10.1038/nature19325

Widloski J, Foster DJ. 2022. Flexible rerouting of hippocampal replay sequences around changing barriers in the absence of global place field remapping. *Neuron* **110**:1547-1558.e8. doi:10.1016/j.neuron.2022.02.002

Wu X, Foster DJ. 2014. Hippocampal replay captures the unique topological structure of a novel environment. *Journal of Neuroscience* **34**:6459–69. doi:10.1523/JNEUROSCI.3414-13.2014