eLife

Reviewed Preprint
v2 • January 7, 2025
Revised by authors

Reviewed Preprint
v1 • August 27, 2024

Neuroscience

Spatial localization of hippocampal replay requires dopamine signaling

Matthew R Kleinman, **David J Foster**

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

eLife Assessment

This work provides a **valuable** contribution to our understanding of the neurobiological mechanisms underlying spatial memory and learning, suggesting that dopamine plays a pivotal role in linking reward context and novelty to memory consolidation processes. The evidence presented to support the main conclusions is **solid**, although reviewers felt that the strength of evidence could have been further strengthened by more rigorous histological verification of the experimental conditions and the complexity of the experimental manipulations, increased sample sizes, and a more consistent approach to experimental dosing and timing, which will be crucial for confirming the reproducibility and reliability of the observed effects.

https://doi.org/10.7554/eLife.99678.2.sa4

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 male 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**, **Figure 1 – supplement 1**), enabling specific and reversible inactivation of VTA dopamine signaling. Post-experiment histology confirmed overlapping virus expression and TH-positive neurons in putative VTA near the injection site (−5.6 mm AP from bregma), as well as approximately 0.5 mm anterior and posterior (−5 to −6 mm AP). Recording microdrives containing 6-32 independently adjustable tetrodes (bilateral 32 tetrode, n=4; unilateral 20 tetrode, n=2; unilateral 6 tetrode, n=1) were implanted above dCA1 (**Fig. 1A**). Tetrodes were lowered to the pyramidal cell layer of dCA1, using the presence of sharp-wave ripples with upward deflections in the LFP, recording depth characteristic of dCA1, and spatially-restricted firing of place cells to confirm the recording location.

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 this reduction was present in both saline and CNO sessions (**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**). Changes in stopping period duration in Epoch 3 were similar across all conditions: Unch. visit duration increased from Epoch 2 to Epoch 3 (**Fig. 1 – Supplement 2E**), while Incr. visit duration decreased (**Fig. 1 – Supplement 2F**). 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. 1 – Supplement 2A-D**). However, the main effect of CNO in experimental rats of prolonging stopping periods occurred in novel sessions (**Fig. 1 – Supplement 2A-B**), not in familiar sessions (**Fig. 1 – Supplement 2C-D**). Rats consistently ran slightly faster towards the Incr. end than the Unch. end in Epoch 2, across all conditions (**Fig. 1 – Supplement 3**).

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.

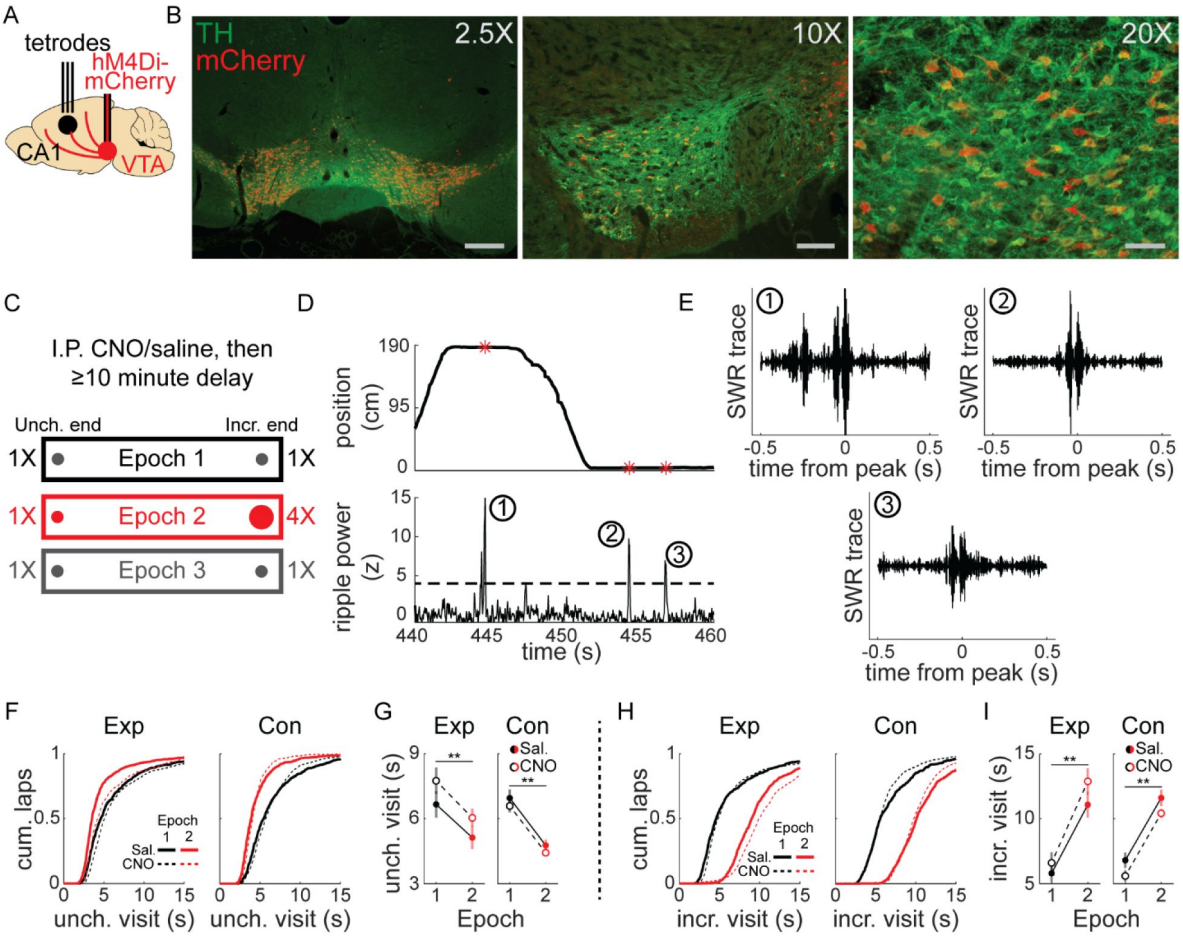


Figure 1.

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 **Figure 1 – Supplement 2**. **(G)** The duration of Unch. reward end stopping periods decreased from Epoch 1 to Epoch 2. The mean stopping durations in each condition were calculated per session, then analyses performed across sessions. Mean ± standard error, Exp Saline, Epoch 1: 6.67±0.61, Epoch 2: 5.13±0.53; Exp CNO, Epoch 1: 7.766±0.61, Epoch 2: 6.04±0.42. Con Saline, Epoch 1: 6.96±0.4, Epoch 2: 4.77±0.27; Con CNO, Epoch 1: 6.59±0.29, Epoch 2: 4.43±0.15. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=-3.37, p<0.001; all other terms, n.s. **(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. The mean stopping durations in each condition were calculated per session, then analyses performed across sessions. Mean ± standard error, Exp Saline, Epoch 1: 6.64±0.54, Epoch 2: 10.87±0.8; Exp CNO, Epoch 1: 7.28±0.66, Epoch 2: 12.31±0.82. Con Saline, Epoch 1: 7.44±0.78, Epoch 2: 11.29±0.5; Con CNO, Epoch 1: 6.47±0.35, Epoch 2: 10.34±0.27. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=4.62, $p<10^{-5}$; all other terms, n.s. *, p<0.05; **, p<0.01.

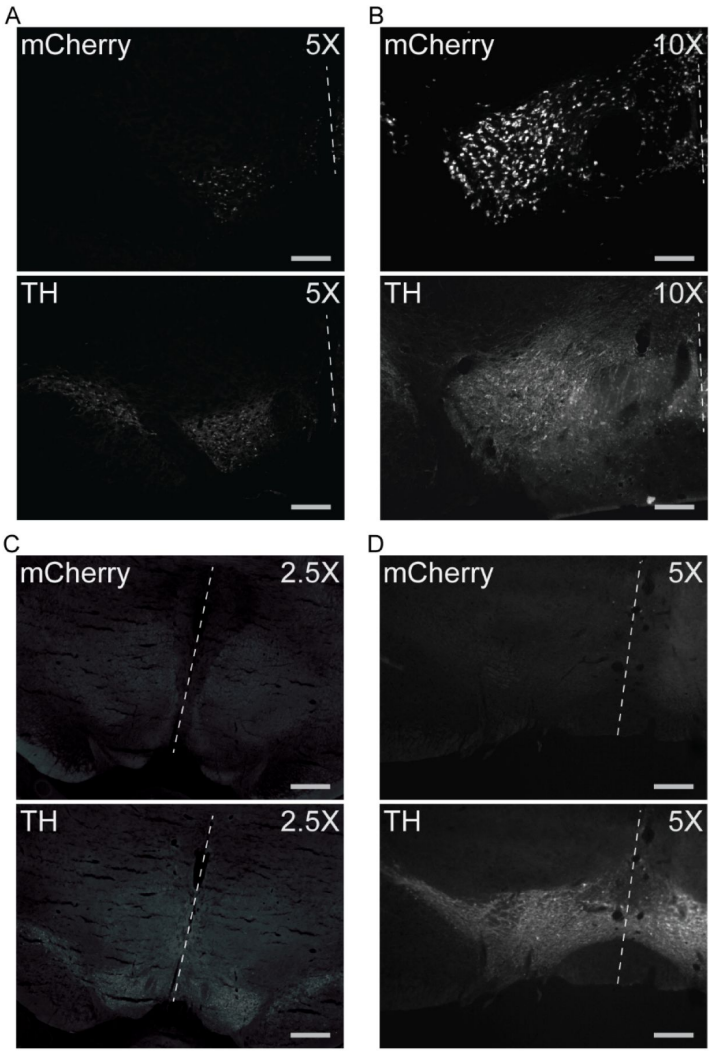


Figure 1 – Supplement 1.

Additional histology examples.

(A) Additional example section from experimental rat with evident virus expression. Scale bar 300 µm. Dashed line marks approximate location of midline. mCherry puncta (top panel) and TH-positive dopaminergic neurons (bottom panel) indicated virus expression in the VTA. **(B)** Additional example section from control rat with evident virus expression. Scale bar 150 µm. **(C)** Example sections from 1st rat excluded due to lack of virus expression. Scale bar 600 µm. Complete lack of mCherry puncta (top panel) indicated failure of virus expression. **(D)** Example sections from 2nd rat excluded due to lack of virus expression. Scale bar 300 µm.

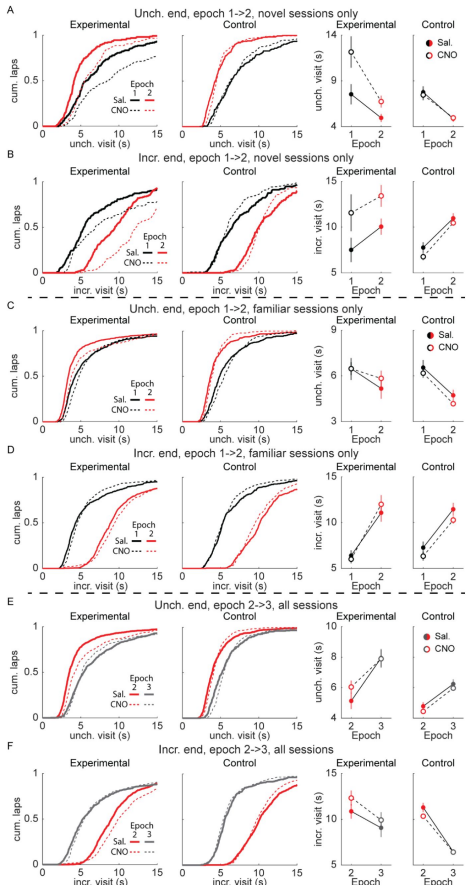


Figure 1 – Supplement 2.

Behavioral effects of novelty and VTA inactivation.

(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: 7.551±1.11, Epoch 2: 4.98±0.44; Exp CNO, Epoch 1: 12.14±1.7, Epoch 2: 6.75±0.64. Con Saline, Epoch 1: 7.76±0.65, Epoch 2: 4.86±0.35; Con CNO, Epoch 1: 7.46±0.59, Epoch 2: 4.96±0.25. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=-2.99, p<0.01; group X drug, z=3.29, p<0.01; all other terms, n.s. **(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: 7.52±1.32, Epoch 2: 10.05±0.88; Exp CNO, Epoch 1: 11.55±2, Epoch 2: 13.39±1.18. Con Saline, Epoch 1: 7.77±0.58, Epoch 2: 10.97±0.55; Con CNO, Epoch 1: 6.77±0.38, Epoch 2: 10.46±0.22. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=2.82, p<0.01; group X drug, z=2.91, p<0.01; all other terms, n.s. **(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: 6.45±0.72, Epoch 2: 5.17±0.66; Exp CNO, Epoch 1: 6.47±0.42, Epoch 2: 5.841±0.51. Con Saline, Epoch 1: 6.54±0.5, Epoch 2: 4.72±0.37; Con CNO, Epoch 1: 6.18±0. 3, Epoch 2: 4.17±0.17. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=-2.37, p<0.05; all other terms, n.s. **(D)** In familiar sessions, Incr. visit duration increased from Epoch 1 to Epoch 2. Mean ± standard error, Exp Saline, Epoch 1: 6.41±0.59, Epoch 2: 11.1±0.97; Exp CNO, Epoch 1: 6.03±0.18, Epoch 2: 11.99±1. Con Saline, Epoch 1: 7.27±0.66, Epoch 2: 11.46±0.71; Con CNO, Epoch 1: 6.33±0.48, Epoch 2: 10.28±0.4. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=3.97, p<0.001; all other terms, n.s. **(E)** Unch. visit duration increased from Epoch 2 to Epoch 3. Mean ± standard error, Exp Saline, Epoch 2: 5.13±0.53, Epoch 3: 7.95±0.6; Exp CNO, Epoch 2: 6.04±0.42, Epoch 3: 7.89±0.68. Con Saline, Epoch 2: 4.77±0.27, Epoch 3: 6.23±0.37; Con CNO, Epoch 2: 4.43±0.15, Epoch 3: 5.96±0.19. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=2.32, p<0.05; all other terms, n.s. **(F)** Incr. visit duration decreased from Epoch 2 to Epoch 3. Mean ± standard error, Exp Saline, Epoch 2: 10.87±0.8, Epoch 3: 9.08±1.01; Exp CNO, Epoch 2: 12.31±0.82, Epoch 3: 9.92±0.95. Con Saline, Epoch 2: 11.29±0.5, Epoch 3: 6.48±0.28; Con CNO, Epoch 2: 10.34±0.27, Epoch 3: 6.42±0.29. Mixed-effects GLM with epoch, drug, animal group, and all interactions, with individual animal as a random effect: epoch, z=-4.77, $p<10^{-5}$, group X epoch, z=2.31, p<0.05; all other terms, n.s.

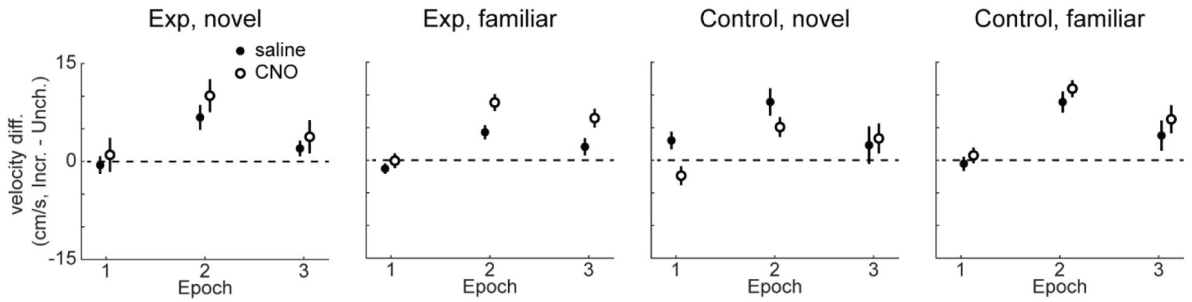


Figure 1 – Supplement 3.

Effect of reward change on running velocity.

Running speed towards the Incr. end in Epoch 2 was consistently significantly faster than towards the Unch. end, across all conditions. The median running speed in all non-zero velocity timepoints while the animal was located outside of the reward end zones in each epoch and running direction was calculated for each session. Mean and standard error across sessions are shown here. Mixed-effects model predicting the velocity difference in Epoch 2, Incr. – Unch., as a function of drug, novelty, and their interaction, with animal-specific intercepts. Experimental group : intercept, $z=2.99$, $p<0.01$; drug, $z=3.18$, $p<0.01$; all other terms, n.s. Control group: intercept, $z=7.85$, $p<10^{-10}$; all other terms, n.s. Filled symbol, saline; unfilled symbol, CNO.

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

We analyzed the rate of SWR occurrence (number of SWR / total duration rat was stationary) 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 ($z=3.19$, $p<0.01$), the interaction between Incr. end and Epoch 2 ($z=9.02$, $p<10^{-10}$) and the three-way interaction between Incr. end, Epoch 2, and CNO ($z=-2.06$, $p<0.05$). 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 ($z=-2.42$, $p<0.05$) and the interaction between Incr. end and Epoch 2 ($z=7.64$, $p<10^{-10}$). SWR duration was reduced in novel sessions, consistent with replays becoming longer with increased familiarity (Berners-Lee et al., 2021), as well as in Epoch 2, but was otherwise unaffected by reward or drug (**Fig. 2 – Supplement 4**).

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. 2 – Supplement 1A**) in novel sessions for both the Unch. end (one-tailed test, CNO>saline, $p<0.001$) and the Incr. end (one-tailed test, saline>CNO, $p<0.01$), as well as at the Unch. end in familiar sessions (one-tailed test, CNO>saline, $p<0.01$). There was no significant difference between saline and CNO in control animals at either reward end in either familiar or novel sessions (**Fig. 2 – Supplement 1B**; one-tailed tests, all n.s.).

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 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; experimental group, full mixed-effects model with drug terms strongly outperformed reduced model lacking drug terms, $AIC_{reduced}–AIC_{full}=5.22$; control group, reduced model lacking drug terms strongly outperformed the full model, $AIC_{reduced}–AIC_{full}=-3.54$). We saw little difference in this measure when comparing sessions that occurred first in each day to those that occurred later (**Fig. 2 – Supplement 3**), and little behavioral evidence from running velocity or number of laps completed that motivation systematically varied across session number. To control for the possibility that VTA inactivation caused changes in locomotor or other non-consummatory

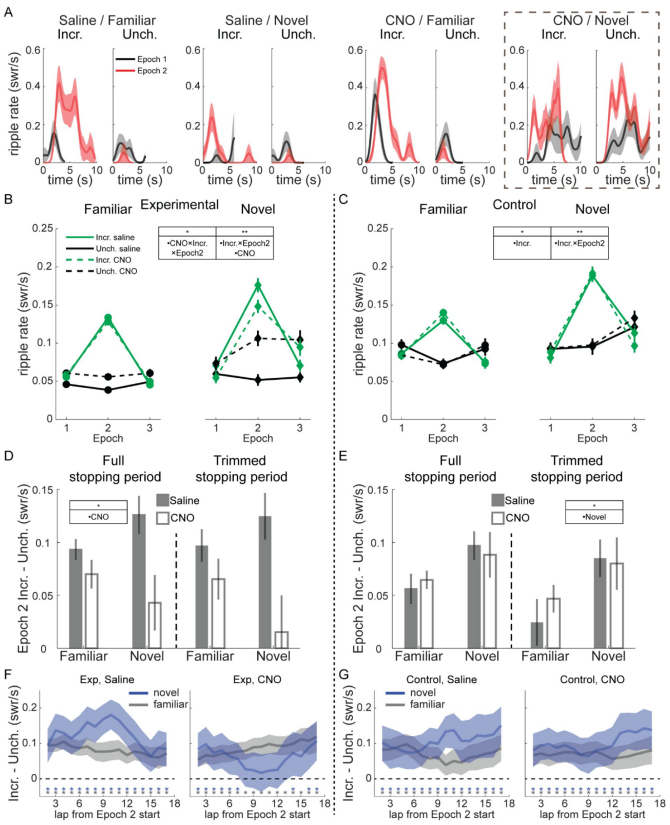


Figure 2.

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). A mixed-effects Poisson generalized linear model (GLM) was fit to predict changes in SWR rate across reward end, epoch, drug condition, and novelty, with animal identity as a random effect. Significant coefficients: CNO (z=3.19, p<0.01), the two-way interaction Incr. end × Epoch 2 (z=9.02, $p<10^{-10}$), and the three-way interaction between Incr. end × Epoch 2 × CNO (z=-2.06, p<0.05). *, p<0.05; **, p<0.01. **(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). The same Poisson GLM fit to control rat data had significant coefficients: Incr. end (z=-2.42, p<0.05) and the two-way interaction Incr. end × Epoch 2 (z=7.64, $p<10^{-10}$). **(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. A mixed-effects model with drug, novelty, and their interaction, and animal-specific intercepts ("full model") was compared to reduced model lacking the drug terms ("reduced model"). Full stopping periods: full model, intercept, z=6.05, $p<10^{-5}$; other terms, n.s. $AIC_{reduced}-AIC_{full}$=5.22. Trimmed stopping periods: full model, intercept, z=3.5, p<0.001; other terms, n.s. $AIC_{reduced}-AIC_{full}$=4.55. **(E)** Difference between SWR rate at Incr. and Unch. ends in Epoch 2 in Control rats, as in (D). Full stopping periods: full model, intercept, z=3.95, $p<10^{-5}$; other terms, n.s. $AIC_{reduced}-AIC_{full}$=-3.54. Trimmed stopping periods: full model, novelty, z=2.31, p<0.05; other terms, n.s. $AIC_{reduced}-AIC_{full}$=-3.1. **(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, uncorrected for multiple comparisons). Shading denotes 95% confidence interval. See also Figure S4. **(G)** As in (F), but for control animals.

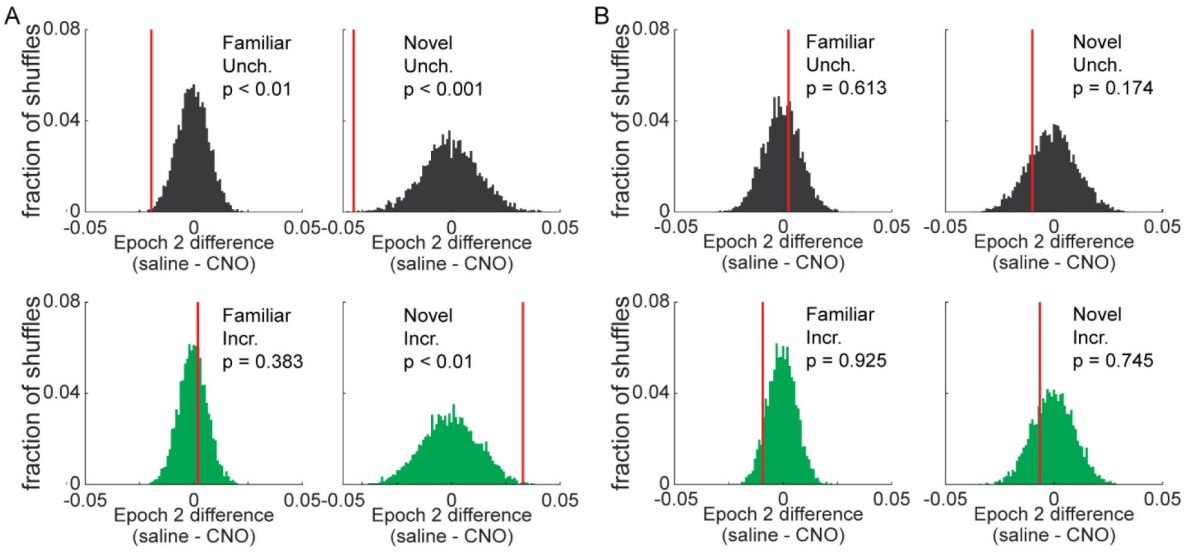


Figure 2 – Supplement 1.

Modulation of SWR rate by reward increase.

(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 2 – Supplement 2.

SWR rate in Epoch 3.

(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.

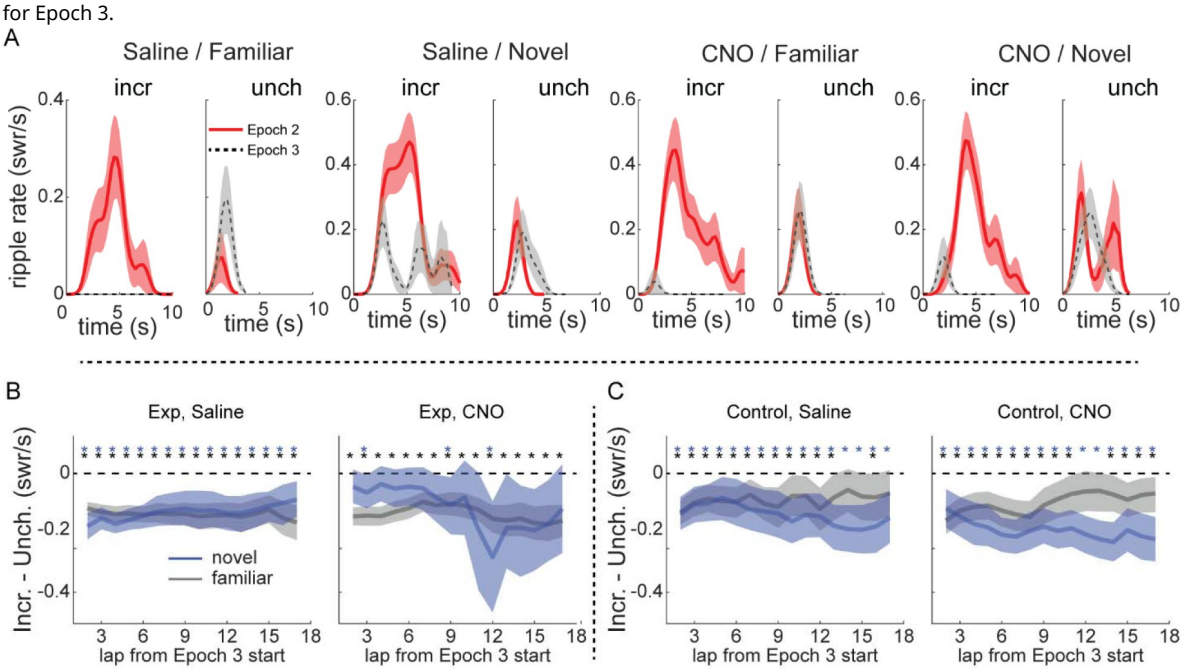
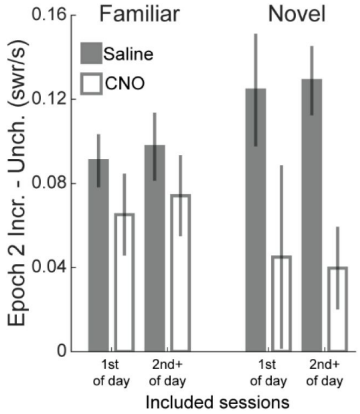


Figure 2 – Supplement 3.

Similar results independent of session number of the day.

The dataset for Experiment 1 was split into two, with one part including only sessions that occurred 1st in any given day ("1st of day") and the other including all sessions that were not the 1st in any given day ("2nd+ of day"). Although the low resultant session counts in each group precluded statistical analysis, the main effect of CNO in reducing SWR rate difference between the reward ends was very similar.



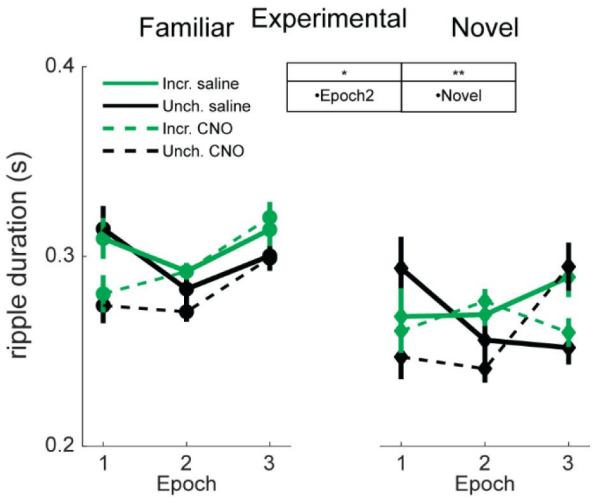


Figure 2 – Supplement 4.

Ripple duration is increased with familiarity.

The duration of SWR was examined similarly to the analysis on SWR rate. A mixed-effects Poisson generalized linear model (GLM) was fit to predict changes in SWR duration across reward end, epoch, drug condition, and novelty, with animal identity as a random effect. Significant coefficients: CNO ($z=-5.86$, $p<10^{-5}$) and Epoch 2 ($z=-2.1$, $p<0.05$).

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; experimental group, full mixed-effects model with drug terms strongly outperformed reduced model lacking drug terms, $AIC_{reduced}$–AIC_{full}=4.55; control group, reduced model lacking drug terms strongly outperformed the full model, $AIC_{reduced}$– AIC_{full}=-3.1).

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, $p<0.05$, uncorrected for multiple comparisons) for almost every task condition and timepoint except in novel CNO sessions in experimental rats (**Fig. 2 – Supplement 2B-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

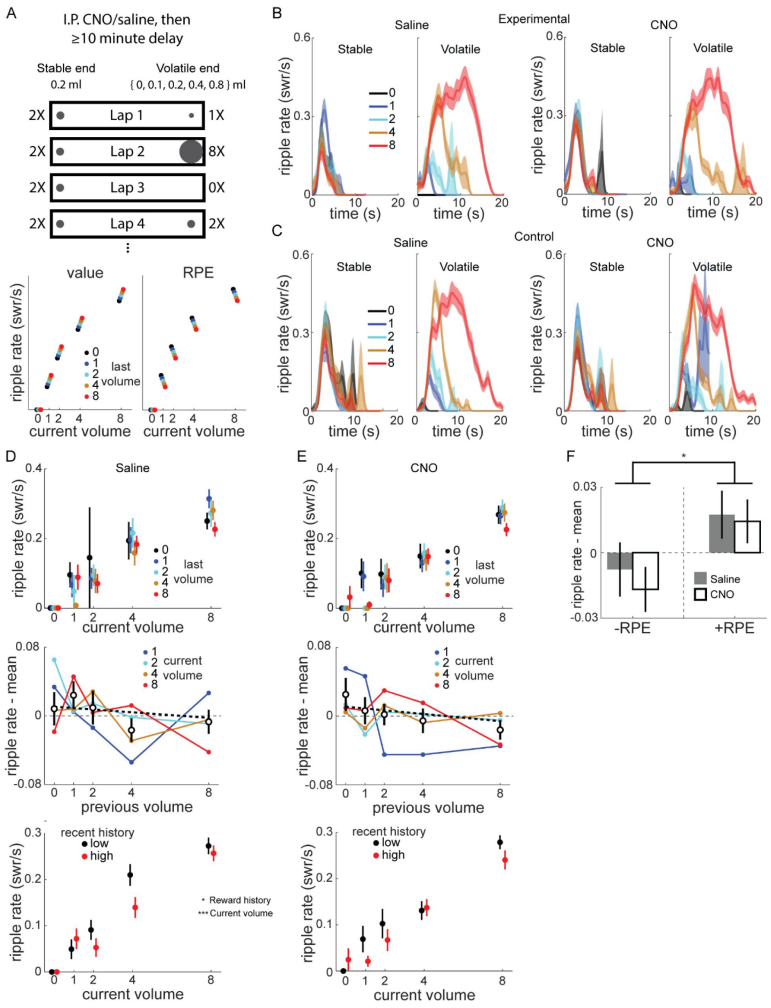


Figure 3.

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 3 – Supplement 1** and **2**. **(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<0.05. GLM fitting SWR rate as a function of drug, current volume, and previous volume: previous volume, z=-2.31, p<0.05; 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), with animal-specific intercept as random effect: volume, z=13.86, $p<10^{-10}$; history, z=-2.23, p<0.05; 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<0.05), drug X RPE sign (F[1,518]=0.07, p=0.785).

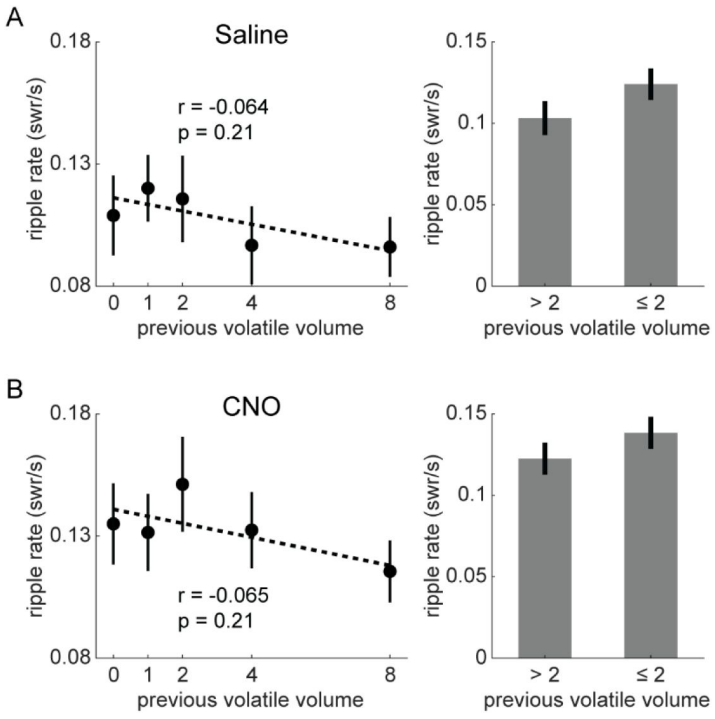


Figure 3 – Supplement 1.

SWR rate at stable end in experimental rats.

(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<0.05), 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.

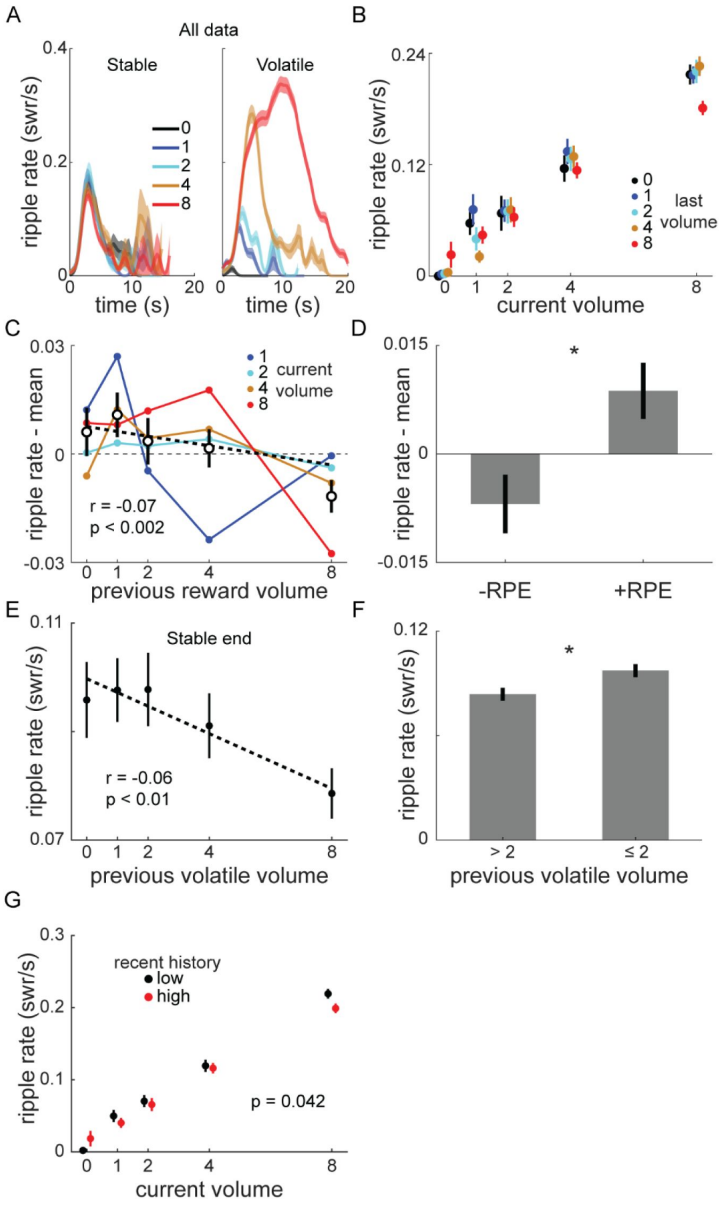


Figure 3 – Supplement 2.

SWR rate in all sessions in volatile reward task.

(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<0.01, 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<0.01). **(E)** SWR rate at the stable end was significantly negatively correlated with the most recent volatile volume (r=-0.06, p<0.01). **(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<0.05).

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. 3 – Supplement 1**).

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 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<0.05). A GLM found the mean-subtracted SWR rate was significantly affected by previous volume (z=-2.3, p<0.05), 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<0.05), 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<0.05), 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. 3 – Supplement 2**). On top of a large increase in SWR rate with current volume at the volatile end (**Fig. 3 – Supplement 2A-B**), SWR rate was also significantly negatively correlated with the previous volatile volume, both at the volatile end (**Fig. 3 – Supplement 2C**) and at the stable end (**Fig. 3 – Supplement 2E**). Accordingly, SWR rates were significantly lower for negative RPE than for positive/non-negative RPE (**Fig. 3 – Supplement 2D-F**). Finally, recent reward history at the volatile end significantly affected SWR rate (**Fig. 3 – Supplement 2G**), 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. 4 – Supplement 1A**). Field similarity across running directions was slightly but significantly increased by both CNO and novelty (**Fig. 4 – Supplement 1B**).

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 (cell count, mean±sem; included sessions: 26.1±1.1, excluded sessions: 9.5±1.6; two-sample t-test, $t(133)=5.3$, $p<10^{-5}$) and also marginally smaller average place field size (mean place field size, defined as number of spatial bins with >1 hz firing rate, mean±sem: included sessions: 47.7±1.3, excluded sessions: 57.7±5.8; two-sample t-test, $t(133)=-2.33$, $p<0.05$), were included for replay analysis (**Fig. 4 – Supplement 2**; 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 (Ambrose et al., 2016) (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 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.

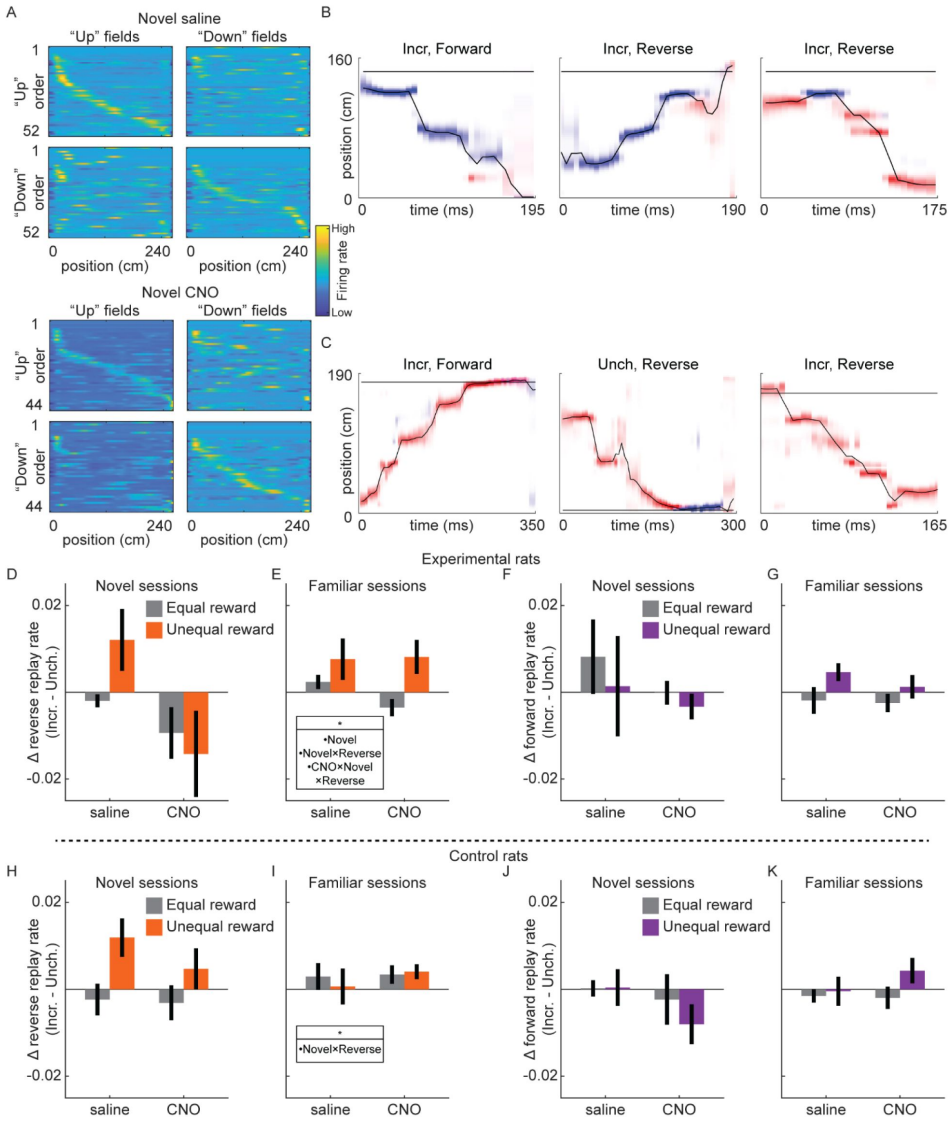


Figure 4.

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). Color denotes z-scored firing rate, from high (yellow) to low (blue). Example saline session and CNO session from experimental rat 3. See also **Figure 4 – Supplement 1** and **2**. **(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 in replay rate between equal and unequal reward conditions was assessed with a mixed-effects linear model with drug, novelty, and replay directionality, and animal-specific intercepts as random effect: drug X novelty X directionality, z=-2.27, p<0.05; novelty, z=-2.01, p<0.05; novelty X directionality, z=2.15, p<0.05; all other terms, n.s. **(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 in replay rate between equal and unequal reward conditions was assessed with a mixed-effects linear model with drug, novelty, and replay directionality, and animal-specific intercepts as random effect: novelty X directionality, z=2.18, p<0.05; all other terms, n.s. **(I)** Same as (H), but for familiar sessions. **(J)** Same as (H), but for forward replay. **(K)** Same as (J), but for familiar sessions.

Figure 4 – Supplement 1.

Effect of novelty and VTA inactivation on place cell properties.

(A) Correlation between single lap place fields and session averaged field. Three-way ANOVA with drug, novelty, and animal group: novelty ($F[1,3249]=6.75$, $p<0.01$), novelty X group ($F[1,3249]=15.76$, $p<0.01$), 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 ($F[1,2816]=5.76$, $p<0.05$), novelty ($F[1,2816]=28.21$, $p<10^{-10}$), drug X novelty ($F[1,2816]=5.52$, $p<0.05$), novelty X group ($F[1,2816]=6.56$, $p<0.05$), all others, $p>0.17$.

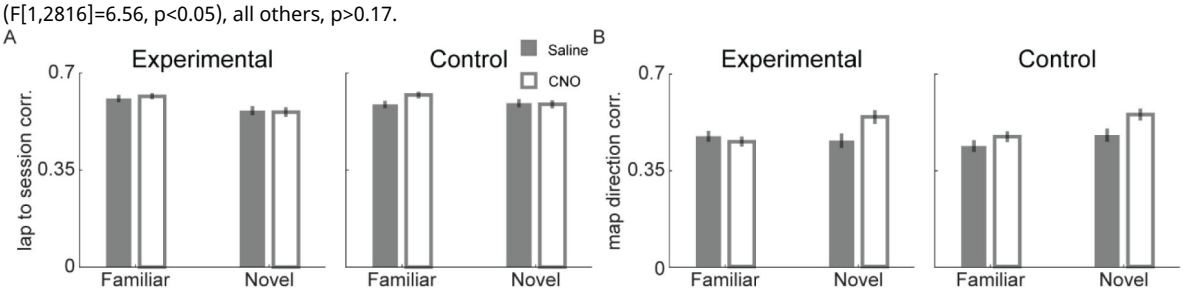
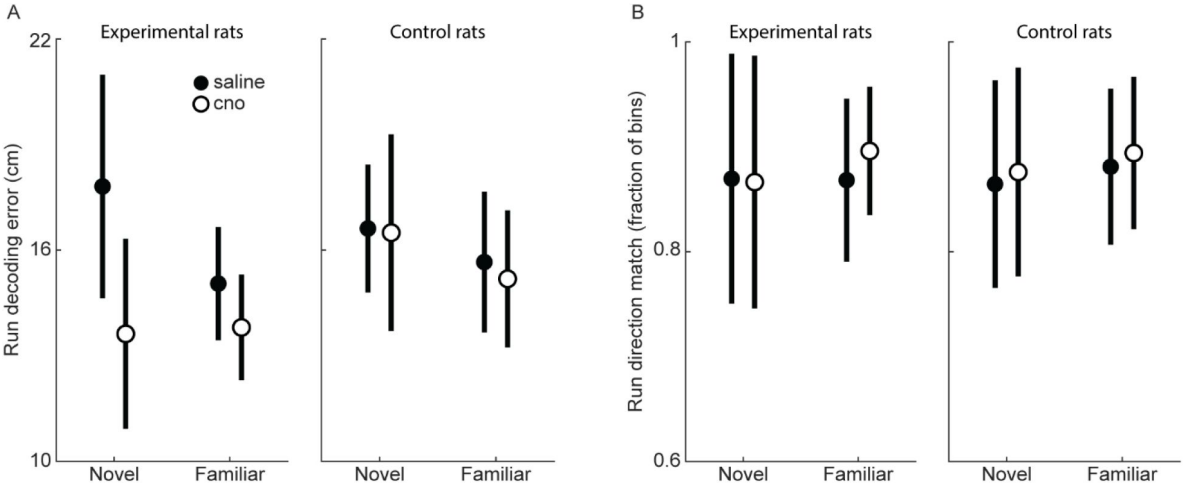


Figure 4 – Supplement 2.

Run decoding accuracy in replay analysis sessions.

(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. A mixed-effects model predicting position decoding error as a function of drug, novelty, recorded neuron count, and mean place field size, found more recorded cells and smaller mean field size in both groups led to smaller decoding errors. Experimental rats: drug, $z=-2.48$, $p<0.05$; cell count, $z=-3.48$, $p<0.01$; mean field size, $z=5.36$, $p<10^{-5}$; other terms, n.s. Control rats: cell count, $z=-4.07$, $p<0.01$; mean field size, $z=4.04$, $p<0.01$; other terms, n.s. Filled and unfilled symbols are saline and CNO sessions, respectively. Error bars, standard error. **(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). A mixed-effects model predicting probability of matching real and decoded run direction as a function of drug, novelty, recorded neuron count, and mean place field size, found more recorded cells in both groups led to higher match probability. Experimental rats: drug, $z=2.33$, $p<0.05$; novelty, $z=-2.96$, $p<0.01$; cell count, $z=6.81$, $p<10^{-5}$; other terms, n.s. Control rats: cell count, $z=6.23$, $p<10^{-5}$; other terms, n.s.



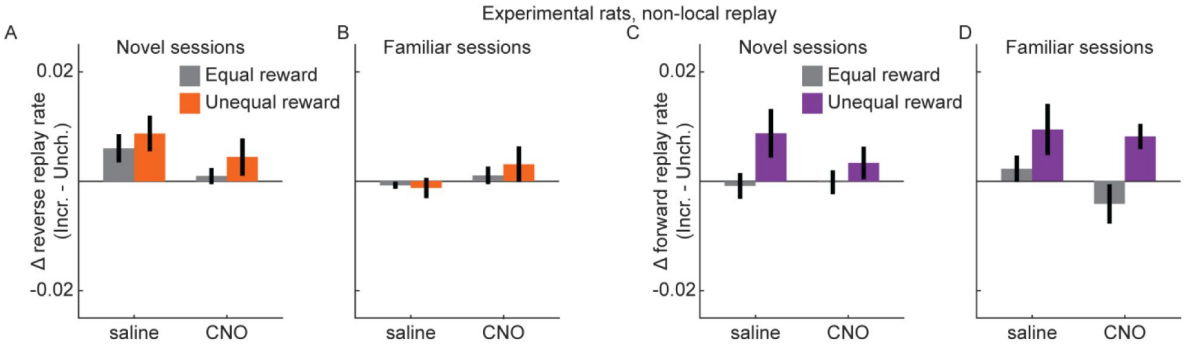


Figure 4 – Supplement 3.

Non-local replay was unaffected by experimental manipulations.

(A) 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 in replay rate between equal and unequal reward conditions was assessed with a mixed-effects linear model with drug, novelty, and replay directionality, and animal-specific intercepts as random effect: all terms, n.s. **(B)** Same as (A), but for familiar sessions. **(C)** Same as (A), but for forward replay. Unequal reward, epoch 2, purple. **(D)** Same as (C), but for familiar sessions.

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 (mixed-effects model with drug, novelty, and replay directionality: novelty X directionality, z=2.18, p<0.05; all other terms, n.s.). The lack of effect of reward in familiar sessions in control rats (**Figure 4I**) was likely due to a dearth of replay in these sessions, with mean reverse replay rates at both reward ends under 0.008 replays/s during Epoch 2, compared to greater than 0.013 replays/s at the Incr. end in Epoch 2 in novel sessions. Reward changes caused no consistent effects in the rates of forward replay (**Fig. 4F-G**, **Fig. 4J-K**), nor in the rates of replay that began far from the animal (non-local replay, **Fig. 4 – Supplement 3**).

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<0.05), 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, n.s.), 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 (mixed-effects model with drug, novelty, and replay directionality: drug X novelty X directionality, z=-2.27, p<0.05; novelty, z=-2.01, p<0.05; novelty X directionality, z=2.15, p<0.05; all other terms, n.s). 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.

Our work has several limitations. First, we did not directly measure the suppression of VTA neurons after CNO injection. Previous work in other brain areas found hM4Di activation suppressed firing rates to around 60% of baseline (Mahler et al., 2014; Chang et al., 2015), in addition to diminished synaptic transmission even when spikes occurred (Stachniak et al., 2014). Combined with the incomplete expression of hM4Di in TH-positive neurons in our animals, we expect VTA activity was significantly but not completely suppressed. Because our results depend only on any degree of blunting differences in dopamine release at different reward locations, rather than the total absence of dopamine signaling, measuring the magnitude of suppression was not essential for our conclusions. Second, we chose simple behavioral tasks specifically to isolate the effect of reward changes per se on hippocampal replay, distinct from learning or performance effects. In more complex tasks, replay overrepresents high value

locations (Pfeiffer and Foster, 2013; Widloski and Foster, 2022), which may aid in consolidating memory of and planning routes to those locations. In our hands, VTA inactivation reduced the SWR rate difference between high and low value locations in novel environments by about 2/3 (**Figure 2D**), which would greatly reduce any effect of reward-biased replay in learning more complex environments or tasks. Finally, we found no effect of VTA inactivation on SWR rates in familiar environments even with large and frequent reward changes (**Figure 3**), but replicating this result with more animals and probing the interaction between volatile reward changes and environmental novelty are important future directions.

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.

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.

Additional information

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.

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.

Drug delivery

At least 10 minutes prior to beginning a recording session (individual animal means of 12 to 17.4 minutes, 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). In Experiment 1 (described below), 2 mg/kg CNO was delivered in all control animal CNO sessions and in most CNO sessions (22/30 sessions), including all novel track CNO sessions, in 3 out of 4 experimental animals. Experimental

rat 3 received 1 mg/kg CNO in all sessions, motivated by our desire to have enough low dosage data in one animal in case 2 mg/kg CNO had an effect in control animals (Gomez et al., 2017). However, this was not the case, and we saw no difference between CNO doses, consistent with some previous reports in the literature (Mahler et al., 2014; Vazey and Aston-Jones, 2014), and therefore combined all data. In Experiment 2, animals received 1mg/kg (n=6 sessions), 2 mg/kg (n=14 sessions), or 4 mg/kg (n=5 sessions), with the higher dose included to probe whether it could reveal a role for VTA in the absence of novelty. As in Experiment 1, we saw little difference between doses and combined all data. 1-4 sessions (mean±s.d., 1.5±0.65) were completed each day, with at least 1.5 hours between injections. Altogether, we were guided in the injection timing and drug doses by previous work in the field (reviewed in Smith et al., 2016), with drug delivery 0-30 minutes before behavioral testing (Kane et al., 2017), significant neuronal suppression occurring within 5-10 minutes and returning to baseline 70 minutes post-injection (Chang et al., 2015; all recording sessions were completed by 60 minutes post injection), and CNO doses between 1 and 10 mg/kg (Smith et al., 2016).

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). Besides this restriction, any other ordering of saline and CNO sessions was permitted. In days with two sessions, for example, saline -> saline, saline -> CNO, and CNO -> CNO all occurred. Nevertheless, we attempted to keep the drug X session number of the day close to balanced. In particular, in Experiment 1 novel sessions (described below), each animal had equal numbers of saline and CNO 1st and 2nd sessions of the day.

Task design

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 1^{st} or 2^{nd} experience on a particular linear track) or “familiar” (3^{rd} 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, all analyses considered SWR rate only during the first 10 s of each end visit to isolate the reward consumption-related period and exclude occasional longer task-disengaged resting periods. Results shown in **Figure 2** were similar if we extended this to examine up to the first 20 s of each reward visit. 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:

- Mean-subtracted rate (x,y) = rate (x,y) − 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

Because decoding quality relies on sufficient spiking and SWR and population spike density events do not always co-occur, our replay analysis used population spiking to determine candidate events. 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, which totaled 4/64 sessions in experimental rats and 12/72 sessions in control rats.

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 log(duration) for each stopping period accounted for the total time the rat was stationary at each reward well visit, 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 analyses 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.

References

Ambrose RE, Pfeiffer BE, Foster DJ (2016) **Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward** *Neuron* **91**:1124–1136 https://doi.org/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 https://doi.org/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 https://doi.org/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 https://doi.org/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** https://doi.org/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 https://doi.org/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 https://doi.org/10.1002/hipo.22488

Chang C. Y., Esber G. R., Marrero-Garcia Y., Yau H.-J., Bonci A., Schoenbaum G (2015) **Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors** *Nature Neuroscience* **19**:1–8 https://doi.org/10.1038/nn.4191

Diba K, Buzsáki G (2007) **Forward and reverse hippocampal place-cell sequences during ripples** *Nat Neurosci* **10**:1241–2 https://doi.org/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* https://doi.org/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 https://doi.org/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 https://doi.org/10.1146/annurev.neuro.30.051606.094341

Foster DJ (2017) **Replay Comes of Age** *Annu Rev Neurosci* **40**:581–602 https://doi.org/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 https://doi.org/10.1038/nature04587

Fujisawa S, Buzsáki G (2011) **A 4 Hz oscillation adaptively synchronizes prefrontal, VTA, and hippocampal activities** *Neuron* **72**:153–65 https://doi.org/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 https://doi.org/10.1016/S0306-4522(96)00202-3

Gomez J. L. *et al.* (2017) **Chemogenetics revealed: DREADD occupancy and activation via converted clozapine** *Science* **357**:503–507 https://doi.org/10.1126/science.aan2475

Gomperts SN, Kloosterman F, Wilson M a. (2015) **VTA neurons coordinate with the hippocampal reactivation of spatial experience** *Elife* **4**:1–22 https://doi.org/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* https://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 https://doi.org/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 https://doi.org/10.1126/science.1217230

Kane G. A., Vazey E. M., Wilson R. C., Shenhav A., Daw N. D., Aston-Jones G., Cohen J. D (2017) **Increased locus coeruleus tonic activity causes disengagement from a patch-foraging task** *Cognitive, Affective and Behavioral Neuroscience* **17**:1073–1083 https://doi.org/10.3758/s13415-017-0531-y

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 https://doi.org/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** https://doi.org/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 https://doi.org/10.1016/S0896-6273(04)00192-8

Kim HGR *et al.* (2020) **A Unified Framework for Dopamine Signals across Timescales** *Cell* **183**:1600–1616 https://doi.org/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 https://doi.org/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 https://doi.org/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 https://doi.org/10.1016/j.neuron.2005.05.002

Mahler S. v., Vazey E. M., Beckley J. T., Keistler C. R., Mcglinchey E. M., Kaufling J., Wilson S. P., Deisseroth K., Woodward J. J., Aston-Jones G. (2014) **Designer receptors show role for ventral pallidum input to ventral tegmental area in cocaine seeking** *Nature Neuroscience* **17**:577–585 https://doi.org/10.1038/nn.3664

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 https://doi.org/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 https://doi.org/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 https://doi.org/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 https://doi.org/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 https://doi.org/10.1038/nature12112

Rosen ZB, Cheung S, Siegelbaum SA (2015) **Midbrain dopamine neurons bidirectionally regulate CA3-CA1 synaptic drive** *Nat Neurosci* **18**:1763–1771 https://doi.org/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 https://doi.org/10.1126/science.1172545

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

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

Smith K. S., Bucci D. J., Luikart B. W., Mahler S. v. (2016) **Behavioral Neuroscience DREADDs: Use and Application in Behavioral Neuroscience DREADDs: Use and Application in Behavioral Neuroscience** *Behavioral Neuroscience* **130**:137–155 https://doi.org/10.1037/bne0000135

Stachniak T. J., Ghosh A., Sternson S. M (2014) **Chemogenetic Synaptic Silencing of Neural Circuits Localizes a Hypothalamus→Midbrain Pathway for Feeding Behavior** *Neuron* **82**:797–808 https://doi.org/10.1016/j.neuron.2014.04.008

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

Takeuchi T *et al.* (2016) **Locus coeruleus and dopaminergic consolidation of everyday memory** *Nature* :1–18 https://doi.org/10.1038/nature19325

Vazey E. M., Aston-Jones G (2014) **Designer receptor manipulations reveal a role of the locus coeruleus noradrenergic system in isoflurane general anesthesia** *Proceedings of the National Academy of Sciences* **111**:3859–3864 https://doi.org/10.1073/pnas.1310025111

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 https://doi.org/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 https://doi.org/10.1523/JNEUROSCI.3414-13.2014

Author information

Matthew R Kleinman
Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley, Berkeley, USA
ORCID iD: 0009-0002-0221-5577

For correspondence: mattrkleinman@berkeley.edu

David J Foster[2]
Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley, Berkeley, USA

For correspondence: davidfoster@berkeley.edu
[2]Lead contact

Editors

Reviewing Editor
Mihaela Iordanova
Concordia University, Montreal, Canada

Senior Editor
Laura Colgin
University of Texas at Austin, Austin, United States of America

Reviewer #1 (Public review):

This manuscript by Kleinman & Foster investigates the dependence of hippocampal replay on VTA activity. They recorded neural activity from the dorsal CA1 region of the hippocampus while chemogenetically silencing VTA dopamine neurons as rats completed laps on a linear track with reward delivery at each end. Reward amount changed across task epochs within a

session on one end of the track. The authors report that VTA activity is necessary for an increase in sharp-wave rate to remain localized to the feeder that undergoes a change in reward magnitude, an effect that was especially pronounced in a novel environment. They follow up on this result with a second experiment in which reward magnitude varies unpredictably at one end of the linear track and report that changes in sharp-wave rate at the variable location reflect both the amount of reward rats just received there, in addition to a smaller modulation that is reminiscent of reward prediction error coding, in which the previous reward rats received at the variable location affects the magnitude of the subsequent change in sharp-wave rate that occurs on the present visit.

This work is technically innovative, combining neural recordings with chemogenetic inactivation. The question of how VTA activity affects replay in the hippocampus is interesting and important given that much of the work implicating hippocampal replay in memory consolidation and planning comes from reward-motivated behavioral tasks.

Comments on revisions:

Overall, I think the authors have done everything they could to address reviewer concerns, short of collecting more data. The more consistent statistical approach makes the paper easier to read and follow. It's helpful to have more details/rationale for the variability in CNO dose and timing. I think some of the results are still not fully convincing, especially the reward volatility experiment (which the authors also note requires additional validation). Given the small number of rats, the small effect sizes, and the complexity of the experimental manipulations, I still have concerns about whether these effects would hold with larger groups sizes.

https://doi.org/10.7554/eLife.99678.2.sa3

Reviewer #2 (Public review):

(1) Summary
Kleinman and Foster's study investigates the role of dopamine signaling in the ventral tegmental area (VTA) on hippocampal replay and sharp-wave ripples (SWR) in rats exposed to changes in reward magnitude and environmental novelty. The authors utilize chemogenetic silencing techniques to modulate dopamine neuron activity in the VTA while conducting simultaneous electrophysiological recordings from the hippocampal CA1 region. Their findings suggest that VTA dopamine signaling is critical for modulating hippocampal replay in response to changes in reward context and novelty, with specific disruptions observed in replay dynamics when VTA is inhibited, particularly in novel environments.

(2) Strengths
The research addresses a significant gap in our understanding of the neurobiological underpinnings of memory and spatial learning, highlighting the importance of dopamine-mediated processes. The methodological approach is robust, combining chemogenetic silencing with precise electrophysiological measurements, which allows for a detailed examination of the neural circuits involved. The study provides important insights into how hippocampal replay and SWR are influenced by reward prediction errors, as well as the role of dopamine in these processes. Specifically, the authors note that VTA silencing unexpectedly did not prevent increases in ripple activities where reward was increased, but induced significant aberrant increases in environments where reward levels were unchanged, highlighting a novel dependency of hippocampal replay on dopamine and a VTA-independent reward prediction error signal in familiar environments. These findings are critical for understanding the consolidation of episodic memory and the neural basis of learning.

(3) Weaknesses
Despite the strengths in methodology and conceptual framework, the study has several weaknesses that could affect the interpretation of the results. There is a need for more rigorous histological validation to confirm the extent and specificity of viral expression (from all animals ideally), which is crucial for ensuring the accuracy of the findings. Variability in the dosing and timing of chemogenetic interventions could also lead to inconsistencies in the data, suggesting a need for more standardized experimental protocols.

Comments on revisions:

I commend the authors for their work in addressing my and the other reviewers' comments. I think these changes have improved the paper, and no further changes are absolutely necessary.

https://doi.org/10.7554/eLife.99678.2.sa2

Reviewer #3 (Public review):

Summary:

The authors of this work are trying to understand the role dopaminergic terminals coming from VTA have on hippocampal mechanisms of memory consolidation, with emphasis on the replay of hippocampal patterns of activity during periods of consummatory behavior in reward locations. Previous work suggested that replay of relevant spatial trajectories supports reward localization and influences behavior.

The authors then tried to separate two conditions that were known to cause an increase in replay activity - spatial novelty encoding and variation of reward magnitude - and evaluate how these changed when VTA dopamine neurons were inactivated by a chemogenetic tool. They found that the rate of reverse replay (trajectory going away from the goal location) is increased with reward only in novel, but not in familiar environments. Overall this suggests that the VTA dopamine signal is critical during learning of novel locations, but not during explorations of already familiar environments.

Strengths:

The inactivation of VTA projections during goal-oriented behavior and in-vivo analysis of patterns of hippocampal activity during both novelty and reward variability. This work adds to the body of evidence that reverse replay constitutes an important mechanism in learning spatial goal locations. Furthermore, this work also points to the role of VTA in reward prediction error with consequences for spatial navigation and consolidation of spatial memories.

The authors addressed very carefully all the points raised during the revision and I am very pleased with the revised manuscript.

https://doi.org/10.7554/eLife.99678.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Reviewer #1:***
>
> *Chemogenetics validation*

> *Little validation is provided for the chemogenetic manipulations. The authors report that animals were excluded due to lack of expression but do not quantify/document the extent of expression in the animals that were included in the study.*

We thank the reviewer for raising this oversight. We have added additional examples of virus expression in sections from included and excluded animals in Figure 1 – Supplement 1. We also added additional comments on the extent of expression we observed in lines 92-95: "Post-experiment histology confirmed overlapping virus expression and TH-positive neurons in putative VTA near the injection site (-5.6 mm AP from bregma), as well as approximately 0.5 mm anterior and posterior (-5 to -6 mm AP)."

> *There's no independent verification that VTA was actually inhibited by the chemogenetic manipulation besides the experimental effects of interest.*

While we did include animals expressing control virus to control for any effect of CNO administration itself, the reviewer is correct that we did not independently verify VTA neurons were inhibited. We have noted this limitation of the current study on lines 513-522 in the Discussion: "We did not directly measure the suppression of VTA neurons after CNO injection. Previous work in other brain areas found hM4Di activation suppressed firing rates to around 60% of baseline (Mahler et al., 2014; Chang et al., 2015), in addition to diminishing synaptic transmission even when spikes occurred (Stachniak et al., 2014). Combined with the incomplete expression of hM4Di in TH-positive neurons in our animals, we expect VTA activity was significantly but not completely suppressed. Because our results depend only on any degree of blunting differences in dopamine release at different reward locations, rather than the total absence of dopamine signaling, measuring the magnitude of suppression was not essential for our conclusions."

> *The authors report a range of CNO doses. What determined the dose that each rat received? Was it constant for an individual rat? If not, how was the dose determined? The authors may wish to examine whether any of their CNO effects were dependent on dose.*

The reviewer is completely correct that we omitted sufficient information regarding the dosage of CNO used in each animal and each session. We have included more details in the Methods lines 676-694, detailing both the doses and the rationale.

> *The authors tested the same animal multiple times per day with relatively little time between recording sessions. Can they be certain that the effect of CNO wore off between sessions? Might successive CNO injections in the same day have impacted neural activity in the VTA differently? Could the chemogenetic manipulation have grown stronger with each successive injection (or maybe weaker due to something like receptor desensitization)? The authors could test statistically whether the effects of CNO that they report do not depend on the number of CNO injections a rat received over a short period of time.*

We thank the reviewer for bringing up the question of whether the order of sessions had an influence on the efficacy of CNO in inactivating VTA activity. To address this, we split our dataset in Experiment 1 into two based on what number session of the particular day it was: 1st sessions of the day vs. all subsequent sessions (2nd+ session of the day). Then, we examined the difference in sharp-wave ripple rate between the reward ends in Epoch 2, as in Figure 2D of the manuscript. Though the resulting number of sessions in each split of the dataset is too low to draw strong statistical conclusions, particularly for novel sessions, we see little evidence there is any systematic change in the effect of VTA inactivation as a

function of session number in the day. We include this in the revised manuscript as Figure 2 – Supplement 3 and in the Results lines 255-258.

> *Motivational considerations*
>
> *In a similar vein, running multiple sessions per day raises the possibility that rats' motivation was not constant across all data collection time points. The authors could test whether any measures of motivation (laps completed, running speed) changed across the sessions conducted within the same day.*

We thank the reviewer for this suggestion. We examined behavioral measures of motivation across sessions conducted within the same day. First, we calculated how many total laps each animal completed each session as a function of the session number of the day. In individual animals, this ranged from -2.8 to 4.1 laps per additional session number (mean 2.01), with an average total laps per session of 43.2 laps. Second, we calculated the median running velocity per session, across both running directions and all epochs, and checked how it varied across session number of the day. Per additional session in the day, this ranged from -3.6 to 8.6 cm/s difference across animals (mean 2.7 cm/s), with an average running velocity of 34.1 cm/s in total. Taken together, while we found little behavioral evidence of strong motivational changes across session, our animals may have been slightly more motivated in later sessions in the day, which also corresponded to later in the light cycle and closer to the dark cycle. We mention this information in Results lines 255-258, related to Figure 2 – Supplement 3.

> *This is a particularly tricky issue, because my read of the methods is that saline sessions were only conducted as the first session of any recording day, which means there's a session order/time of day and potential motivational confound in comparing saline to CNO sessions.*

We have clarified the ordering of CNO and saline sessions in the Methods lines 697-702. Briefly, we avoided running CNO sessions before saline sessions in the same day, but either could be the first session of a day. That is, saline -> saline, saline -> CNO, and CNO -> CNO were all valid orderings. On days with more than two sessions, any number of repeated saline and CNO sessions was permitted, provided that as soon as a CNO session occurred, any subsequent sessions were also CNO.

More generally, we shared this reviewer's concern about potential confounds between drug and motivation. For novel sessions in Experiment 1, each animal had equal numbers of saline and CNO 1st and 2nd sessions of the day. For familiar sessions, animals had similar counts for 1st sessions of the day (experimental rats: 20 saline, 16 CNO; control rats: 17 saline, 15 CNO) but more CNO 2nd sessions of the day (experimental rats: 5 saline, 13 CNO; control rats: 5 saline, 10 CNO). There were occasionally 3rd or 4th sessions in a given day for some rats, and these were also approximately equal (experimental rat 2, 3rd sessions: 2 each of saline and CNO, 4th session: 1 saline; experimental rat 3 and 4, 3rd sessions: 1 each of saline and CNO; control rat 2, 3rd session: 1 saline).

> *Statistics, statistical power, and effect sizes*
>
> *Throughout the manuscript, the authors employ a mixture of t-tests, ANOVAs, and mixed-effects models. Only the mixed effects models appropriately account for the fact that all of this data involves repeated measurements from the same subject. The t-tests are frequently doubly inappropriate because they both treat repeated measures as independent and are not corrected for multiple comparisons.*

We thank the reviewer for pointing out these issues with our statistical analyses in places. We have made the following improvements:

Figure 1F-I, S1, reward end visit durations: We now use a linear mixed-effects model to analyze the difference in stopping period durations between epochs. For each session, we calculated the mean stopping duration for each reward end in each epoch, then modeled the difference between epochs as a function of drug and novelty, with animal-specific intercepts. For example, related to Figure 1F and also described in the Results, we modeled the stopping duration difference at the Unchanged reward end, Epoch 2 – Epoch 1, and found experimental rats had a significant intercept (Epoch 2 stops shorter than Epoch 1) and the drug × novelty interaction, while control rats had a significant intercept and novelty main effect. The other visit duration analysis shown in Figure 1 – Supplement 1 have similarly been updated.

Figure 2D-E, ripple rate difference between reward ends in Epoch 2: We now use a linear mixed-effects model to analyze the difference between ripple rates at the Incr. and Unch. reward ends in Epoch 2. For each session, we calculated the mean ripple rate at each end in Epoch 2, then modeled the difference as a function of drug and novelty, with animal-specific intercepts. With the full stopping periods, for experimental rats, there was a significant intercept (ripple rate at Incr. greater than Unch.) and the model with drug included performed significantly better than the one without it ($AIC_{nodrug} - AIC_{full} = 5.22$). Control rats had a significant intercept and effect of novelty (greater difference with novelty), and the model excluding drug terms performed better ($AIC_{nodrug} - AIC_{full} = -3.54$). Results with the trimmed stopping periods were similar. These analyses are described in Results lines 253-266.

Figure 3D-E, ripple rate as a function of reward history: We now use a mixed-effects model that incorporates animal-specific intercepts. The results remained similar and have been updated in the text and legend.

Figure 4D-K, replay rates as a function of drug, novelty, and directionality: We now use mixed-effects models that incorporate animal-specific intercepts rather than three-way ANOVA. The results remained similar and have been updated in the text and legend.

> *The number of animals in these studies is on the lower end for this sort of work, raising questions about whether all of these results are statistically reliable and likely to generalize. This is particularly pronounced in the reward volatility experiment, where the number of rats in the experimental group is halved to just two. The results of this experiment are potentially very exciting, but the sample size makes this feel more like pilot data than a finished product.*

We have added additional emphasis in the text that the experimental group results of CNO inactivation in the volatile reward task should be confirmed with future work, in Discussion line 529-533. Because these experiments were performed on familiar tracks, we see them as corroborating/complementing the results from Experiment 1. Although the analysis assumes VTA inactivation had no effect, our pooling of all Experiment 2 data to display in Figure 3 – Supplement 2 maximized our ability to analyze the effects of volatile reward deliveries on sharp-wave ripple rates, lending further support to the main results shown in Figure 3.

> *The effect sizes of the various manipulations appear to be relatively modest, and I wonder if the authors could help readers by contextualizing the magnitude of these results further. For instance, when VTA inactivation increases mis-localization of SWRs to the unchanged end of the track, roughly how many misplaced sharp-waves are occurring within a session, and what would their consequence be? On this particular behavioral task, it's not clear that the animals are doing worse in any way despite the mislocalization of sharp-waves. And it seems like the absolute number of extra sharp-waves that occur in some of these conditions would be quite small over the course of a*

> *session, so it would be helpful if the authors could speculate on how these differences might translate to meaningful changes in processes like consolidation, for instance.*

We thank the reviewer for this helpful suggestion to give some context to the difference in sharp-wave ripple numbers and the functional consequence of these changes. We agree completely that this task is almost certainly too simple for animals to show any performance deficit from these changes. We chose this precisely so we could examine the consequences of VTA inactivation to the sharp-wave ripple response to reward changes per se, without any confound of performance or memory changes that could also conceivably alter sharp-wave ripples. We have added both more context about the magnitude and consequence of these sharp-wave ripple changes as well as comments about the choice of this particular task (Discussion lines 522-529).

> *How directly is reward affecting sharp-wave rate?*
>
> *Changes in reward magnitude on the authors' task cause rats to reallocate how much time they spent at each end. Coincident with this behavioral change, the authors identify changes in the sharp-wave rate, and the assumption is that changing reward is altering the sharp-wave rate. But it also seems possible that by inducing longer pauses, increased reward magnitude is affecting the hippocampal network state and creating an occasion for more sharp-waves to occur. It's possible that any manipulation so altering rats' behavior would similarly affect the sharp-wave rate.*
>
> *For instance, in the volatility experiment, on trials when no reward is given sharp-wave rate looks like it is effectively zero. But this rate is somewhat hard to interpret. If rats hardly stopped moving on trials when no reward was given, and the hippocampus remained in a strong theta network state for the full duration of the rat's visit to the feeder, the lack of sharp-waves might not reflect something about reward processing so much as the fact that the rat's hippocampus didn't have the occasion to emit a sharp-wave. A better way to compute the sharp-wave rate might be to use not the entire visit duration in the denominator, but rather the total amount of time the hippocampus spends in a non-theta state during each visit. Another approach might be to include visit duration as a covariate with reward magnitude in some of the analyses. Increasing reward magnitude seems to increase visit duration, but these probably aren't perfectly correlated, so the authors might gain some leverage by showing that on the rare long visit to a low-reward end sharp-wave rate remains reliably low. This would help exclude the explanation that sharp-wave rate follows increases in reward magnitude simply because longer pauses allow a greater opportunity for the hippocampus to settle into a non-theta state.*

We thank the reviewer for these important comments. We have better clarified the analysis of sharp-wave ripple rate in the Results (lines 172-173). To speak to the main concern of the reviewer, we do only consider times during “stopping periods” when the rat is actually stationary. That is, ripple rate for each visit is calculated as (# of ripples / total stationary time), rather than the full duration the rat is at the track end. With respect to including visit duration as a covariate, the Poisson model takes the total stationary time of each visit into account, so that it is effectively predicting the number of events (ripples) per unit of time (seconds) given the particular experimental variables (reward condition, drug condition, etc.). We have added additional clarification of this in the Methods (line 834-836).

> *The authors seem to acknowledge this issue to some extent, as a few analyses have the moments just after the rat's arrival at a feeder and just before departure trimmed out of consideration. But that assumes these sorts of non-theta states are only occurring at the very beginning and very end of visits when in fact rats might be doing all sorts of other*

> *things during visits that could affect the hippocampus network state and the propensity to observe sharp-waves.*

We hope that with the clarification provided above, this control analysis helps remove any potential effects of approaching/leaving behavior or differences in movement at the reward end that could alter sharp-wave ripple rates.

> *Minor issues*
>
> *The title/abstract should reflect that only male animals were used in this study.*

We have added this important information to the Abstract line 21.

> *The title refers to hippocampal replay, but for much of the paper the authors are measuring sharp-wave rate and not replay directly, so I would favor a more nuanced title.*

We thank the reviewer for this suggestion. In the context of our work, we consider sharp-wave ripples as more-easily-detected markers for the occurrence of replay. Previous work from our lab (Ambrose et al., 2016) showed the effect of reward changes had very similar effects to both sharp-wave ripple rate and replay rate. We try to be explicit about viewing ripples as markers of replay content in both the Introduction and Discussion. Nevertheless, we do also demonstrate the title claim directly – by measuring replay and its spatial localization – therefore we feel comfortable with the title as it is.

> *Relatedly, the interpretation of the mislocalization of sharp-waves following VTA inactivation suggests that the hippocampus is perhaps representing information inappropriately/incorrectly for consolidation, as the increased rate is observed both for a location that has undergone a change in reward and one that has not. However, the authors are measuring replay rate, not replay content. It's entirely possible that the "mislocalized" replays at the unchanged end are, in fact, replaying information about the changed end of the track. A bit more nuance in the discussion of this effect would be helpful.*

While we do show that replay content, in the form of reverse vs. forward replays, is altered with VTA inactivation, we take the reviewers point and completely agree. Especially in the context of the linear track, replays at either end could certainly be updating/consolidating information about both ends. We would argue our results suggest VTA is critical to localizing ripples and replay in more complex environments where this is not the case, but this is a hypothesis. We have added clarification and discussion of this point (Discussion lines 522-529).

However, in response to the reviewer's comment, we have now also examined non-locally-initiated replays specifically to determine whether the increased ripple rate at the Unch. reward end in novel CNO sessions was likely due to more non-local replay, but found no significant increases in non-local replay at either reward end in either drug condition or novelty condition. We have included this result as Figure 4 – Supplement 3, and note it in the Results lines 487-488.

> *The authors use decoding accuracy during movement to determine which sessions should be included for decoding of replay direction. Details on cross-validation are omitted and would be appreciated. Also, the authors assume that sessions failed to meet inclusion criteria because of ensemble size, but this information is not reported anywhere directly. More info on the ensemble size of included/excluded sessions would be helpful.*

We have added additional information about the run decoding procedure and related session inclusion criteria, as well as about recorded ensemble sizes (lines 417-421). Briefly, mean ensemble sizes were significantly smaller for excluded sessions (cell count, mean±sem; included sessions: 26.1±1.1, excluded sessions: 9.5±1.6; two-sample t-test, $t(133)=5.3$, $p<10^{-5}$). The average field size, defined as the number of spatial bins with greater than 1 hz firing rate, in excluded sessions was also larger (mean±sem, included sessions: 47.7±1.3, excluded sessions: 57.7±5.8; two-sample t-test, $t(133)=-2.33$, $p<0.05$), though the difference was less dramatic. Using a mixed effects model to predict position decoding error (as in Figure 4 – Supplement 2A) as a function of drug, novelty, cell count, and mean place field size, in both experimental and control groups cell count and field size were significant predictors: more cells and smaller average field size led to lower error. A similar model that instead predicted the fraction of running bins with correctly decoded running direction (as in Figure 4 – Supplement 2B), in neither group was field size significant, while cell count remained so: more cells led to more bins with running direction correctly classified. We include these analyses in the legend for the figure. With respect to cross validation of run decoding, because both the contribution of spikes in any single time bin to a neuron's place field is extremely small and because we used run decoding accuracy simply to filter out sessions with poorer decoding, we did not use cross validation here.

> *For most of the paper, the authors detect sharp-waves using ripple power in the LFP, but for the analysis of replay direction, they use a different detection procedure based on the population firing rate of recorded neurons. Was there a reason for this switch? It's somewhat difficult to compare reported sharpwave/replay rates of the analyses given that different approaches were used.*

We have added clarification for this change in detecting candidate events (lines 787-789). Briefly, sharp-wave ripples and spike density events are often but not always overlapping, such that there can be strong ripples with little spiking in the recorded ensemble or weak/absent ripples during vigorous spiking in the recorded ensemble. Because the decoding of replay content relies on spiking, our lab and others often use spike density or population burst events as candidate events. We have confirmed that the main results of Experiment 1 (e.g., Figure 2) remain the same if we use spike density events rather than sharp-wave ripples, but prefer to keep the use of sharp-wave ripples here for better comparison with Experiment 2 and to allow the inclusion of animals and sessions with low cell yield but clear ripples in the LFP.

> ***Reviewer #2 (Recommendations For The Authors):***
>
> *Include additional histological data to confirm the extent of viral spread and precise tetrode placements. Providing detailed figures that clearly illustrate these aspects would strengthen the validity of the neural recordings and the specificity of the chemogenetic silencing.*

We thank the reviewer for this suggestion and have added additional information regarding virus expression in Figure 1 – Supplement 1. We also added additional comments on the extent of expression we observed in lines 92-95: "Post-experiment histology confirmed overlapping virus expression and TH-positive neurons in putative VTA near the injection site (-5.6 mm AP from bregma), as well as approximately 0.5 mm anterior and posterior (-5 to -6 mm AP)."

While we do not show histological confirmation of hippocampal recording sites, the presence of sharp-wave ripples with upward deflections, presence of place cells, and recording coordinates and depth typical of dorsal CA1 made us confident in our recording location. We have noted these characteristics of our recordings in lines 128-131 in the Results: "Tetrodes

were lowered to the pyramidal cell layer of dCA1, using the presence of sharp-wave ripples with upward deflections in the LFP, recording depth characteristic of dCA1, and spatially-restricted firing of place cells to confirm the recording location."

> *Address the variability in CNO dosing and timing before recordings. It is recommended to standardize the dose and ensure a consistent timing interval between CNO administration and the start of recordings to minimize variability in the effects observed across different subjects. Instead of collecting new data, the authors could report the data for each animal, indicating the dose and interval between the injection and the recording.*

We have further clarified the CNO dosing and timings in lines 676-702.

> *In Figure 1F, explicitly state whether the data represent averages across multiple sessions and confirm if these observations are primarily from the initial novel sessions. This clarification will help in accurately interpreting the effects of novelty on the measured neural activities.*

We have changed the analyses shown in Figure 1F-I and Figure 1 – Supplement 1 thanks to the suggestions of Reviewer #1, but also more clearly spell out the analysis. Briefly, we average the durations for each condition within session (e.g., take the mean Unch. duration in Epoch 1), then perform the analysis across sessions. These data come from all sessions in Experiment 1, as described in lines 141-147, meaning there are around 2-3 times as many familiar sessions as novel sessions.

> *Reconsider the reporting of marginal p-values (e.g., p=0.055). If the results are borderline significant, either more data should be collected to robustly demonstrate the effects or a statistical discussion should be included to address the implications of these marginal findings.*

We have removed the reporting of marginal p-values.

> *Ensure that the axes and scales are consistent across similar figures (specifically mentioned for Figure 2A) to prevent misinterpretation of the data. Consider showing the average across all animals in 2A, similar to 2B and 2C.*

We have adjusted these axes to be consistent across all panels.

> *Add a legend to the heatmap in Figure 4A to facilitate understanding of the data presented.*

We have added a heatmap to the figure and legend.

> *Provide a detailed examination and discussion of the apparent contradictions observed in control data, particularly where experimental conditions with saline show increased reverse replay in novel environments, which is absent in familiar sessions. See Figures 4E and 4I.*

We thank the reviewer for noting that this feature of our data deserved discussion. We confirmed that the lack of an effect of reward on reverse replay rates in familiar sessions in control rats was due to generally low replay rates in these sessions. Replay rates have been observed to decrease as the familiarity of an environment or behavior increases, and the presence of the reward-related modulation of reverse replay in novel sessions in these animals is consistent with this observation. We now report in the Results lines 458-459 and

485-486 the low replay rates in this group in familiar sessions, and the likelihood that this is preventing any reward-related modulation from being detected.

> *Include a more detailed analysis of place cell properties, such as firing rates and field sizes, especially in novel environments where VTA inactivation appears to alter spatial coding. Decoding error is lower during CNO administration - does this mean place fields are smaller/more accurate? This analysis could offer deeper insights into the mechanisms by which dopamine influences hippocampal neural representations and memory processes.*

We thank the reviewer for this helpful suggestion. We have expanded on our analysis of place field properties and decoding accuracy, describing properties of sessions with good enough decoding to be included compared to those that were excluded (lines 417-421). We also directly tested how decoding quality depended on several factors, including drug condition, novelty, number of cells recorded, and the average place field size of recorded cells (see legend for Figure 4 – Supplement 2). We found a small but significant effect of drug in experimental rats, but larger effects of number of recorded cells and average field size, that were also present in control animals.

> *Correct the typo on line 722 from "In ANOVA" to "An ANOVA".*

We reworded this section and have corrected this error.

> ***Reviewer #3 (Recommendations For The Authors):***
>
> *The manuscript is clear and exciting. As a main criticism, I would have liked to see the effects on ripple duration not just the rate.*

We thank the reviewer for this interesting idea. We performed a new analysis, similar to our analysis on SWR rate, probing the effect of our experimental manipulations on SWR duration in experimental rats. We have added the results in Figure 2 – Supplement 4, and note them in the main text lines 195-198: “SWR duration was reduced in novel sessions, consistent with replays becoming longer with increased familiarity (Berners-Lee et al., 2021), as well as in Epoch 2, but was otherwise unaffected by reward or drug (Figure 2 – Supplement 4).”

> *I have a few other minor comments:*
>
> *(1) I find it a little disturbing and counterintuitive that statistical differences are not always depicted in the figure graphs (for example Figures 2A-E). If the authors don't like to use the traditional *, ** or *** they could either just use one symbol to depict significance or simply depict the actual p values.*

We thank the reviewer for this suggestion. We struggled with indicating significance values graphically in an intuitive way for interaction terms in the figures. We now added significance indicators in Figures 1F-I, added the significant model coefficients directly into Figure 2B-C, changed the analysis depicted in Figure 2D-E per Reviewer 1’s suggestions, and added significance indicators where previously missing in Figures 3 and 4.

> *(2) Related to the point above: in the page 7 legend D and E, it would be advantageous for clarity of the experimental results to also perform post-hoc analyses as depicted in the graphs, rather than just describe the p-value of the 3way ANOVA;*

We thank the reviewer for this suggestion. Because the figure includes the mean and standard error of each condition, in addition to the significant effects of the mixed-effects model, we prefer the current format as it makes clearer the statistical tests that were

performed while still allowing visual appreciation of differences between specific experimental conditions of interest to the reader.

> *(3) According to Figure 1H, the duration of the reward visits can go up to 15s (or more). Yet in Figure 2A only the first 10sec were analyzed. While I understand the rationale for using the initial 10 seconds where there is a lot more data, the results of graphs of Figures A to C will not have the same data/rate as Figures D-F where I assume the entire duration of the visit is taken into account.*
>
> *A figure showing what happening to the ripple rate during the visits >10sec would help interpret the results of Figure 2.*

We thank the reviewer for these interesting suggestions. We clarify now that all these analyses of Experiment 1 use only the first 10 s of each stopping period in Method line 758-764. However, examining the longer stopping periods is an excellent suggestion, and we re-analyzed the Experiment 1 dataset using up to the first 20 s of each stopping period. The main results (e.g., Figure 2) remain the same:

(1) Related to Figure 2B-C: For experimental rats, a mixed-effects generalized linear model predicting sharp-wave ripple rate as a function of reward end, block, drug, novelty, and interactions, had the following significant terms: drug ($p<10^{-5}$), novelty ($p<10^{-10}$), reward end × block ($p<10^{-10}$), and reward end × block × drug ($p<0.05$). The same model in control rats had significant terms: reward end ($p<0.05$), novelty ($p<10^{-4}$), reward end × block ($p<10^{-10}$).

(2) Related to Figure 2D-E: For experimental rats, we used a mixed-effects generalized linear model predicting the difference in sharp-wave ripple rate between the Incr. and Unch. reward ends in Epoch 2 as a function of novelty, drug, and their interaction. Model comparison found the full model performed better than a model removing the drug terms ($AIC_{nodrug} - AIC_{full} = 2.94$), while a model with only the intercept performed even worse ($AIC_{intercept} - AIC_{full} = 13.76$). For control rats, model comparison found the full model was equivalent to a model with only the intercept (AICintercept – AICfull = -0.36), with both modestly better than a model with no drug terms ($AIC_{nodrug} - AIC_{full} = 1.38$).

We have added a remark that results remain the same using this longer time window in Methods line 758-764.

https://doi.org/10.7554/eLife.99678.2.sa0