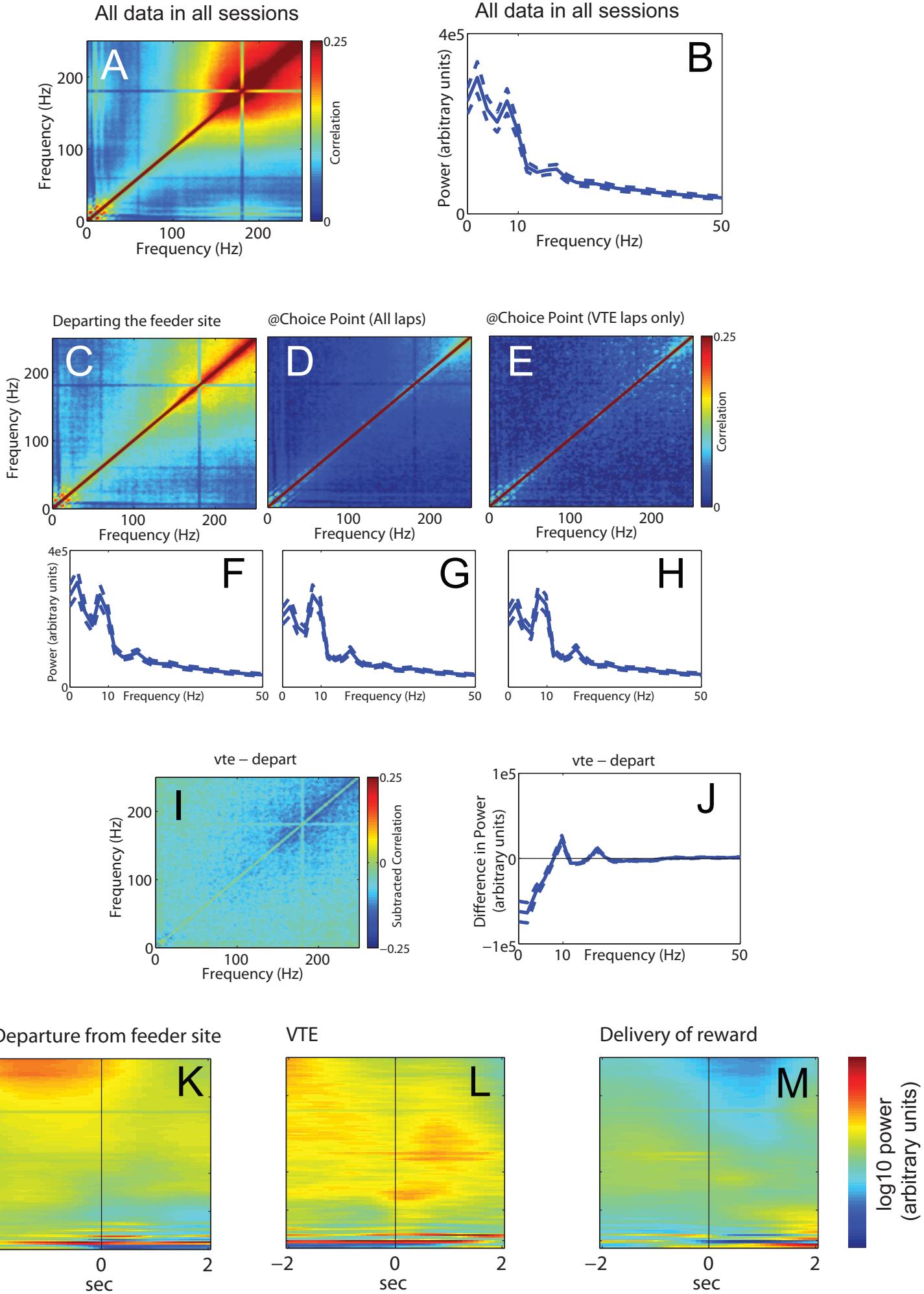
Supplemental Figure 1

Associated with Figures 2 and 3.

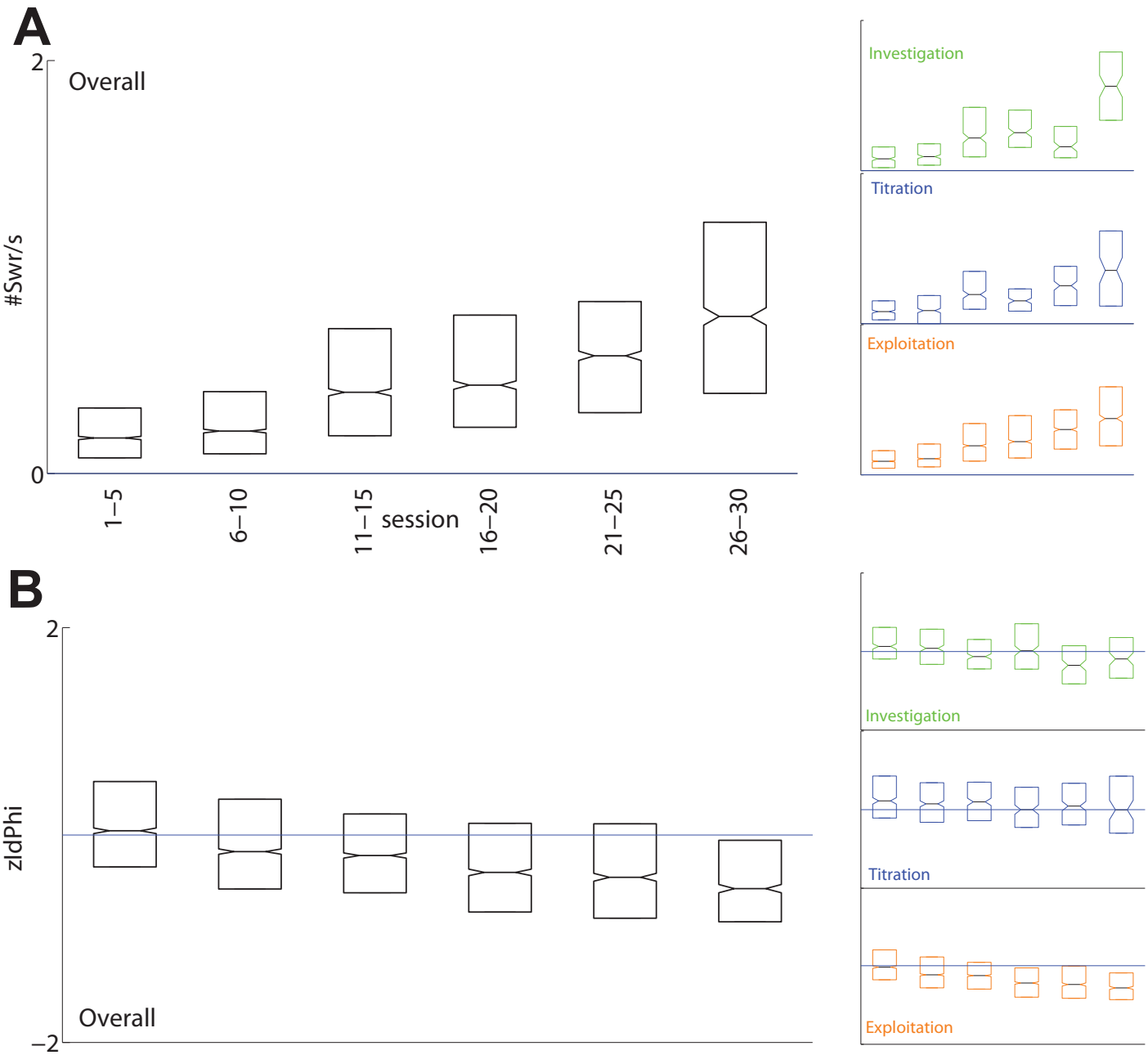
Local Field Potentials differed between departing the reward site and passing through the choice point. A Spectral auto-coherence (SAC) showing increased correlation at sharp wave frequencies (between 150 and 250 Hz). Spectral auto-coherence is calculated as the correlation of the spectrogram, collapsed across time. These auto-coherence plots can be used to identify transient events of consistent frequency. The sharp lines at 60 and 180 Hz are line noise that is uncorrelated to neural signals. **B**: Power spectral density (PSD) showing that there is power in the delta (2-4 Hz) and theta (6-10 Hz) bands. Panels **A** and **B** are taken over all data from all sessions on the DD task. **C**,**F**: SAC and PSD plots for the 3s preceding departure from the feeder-zone. Note the strong SWR events in the auto-coherence plot and the lack of strong theta power ($\theta < \delta$). **D**,**G**: SAC and PSD plots for the 3s preceding leaving the choice point. Note that there is little to no SWR signals in the choice point passes, but there is strong theta power ($\theta > \delta$). **E**,**H**: SAC and PSD plots for the 3s surrounding VTE events. Note that there is little to no SWR signals in the choice point passes, but there is strong theta power ($\theta > \delta$). The difference can be seen clearly in the subtraction plots (**I**,**J**). **K**,**L**,**M**: Frequency-normalized average spectrograms aligned to the departure from the reward site, VTE events, and delivery of reward. Note the gamma frequencies at the VTE event show increased high gamma and high-frequency oscillations as the animals re-orient at the turn around point.


All data in all sessions
A
Frequency (Hz)
Correlation
All data in all sessions
B
Power (arbitrary units)
Frequency (Hz)
Departing the feeder site
@Choice Point (All laps)
@Choice Point (VTE laps only)
C
D
E
F
G
H
vte – depart
I
Subtracted Correlation
vte – depart
J
Difference in Power (arbitrary units)
Departure from feeder site
VTE
Delivery of reward
K
L
M
Hz
sec
log10 power (arbitrary units)

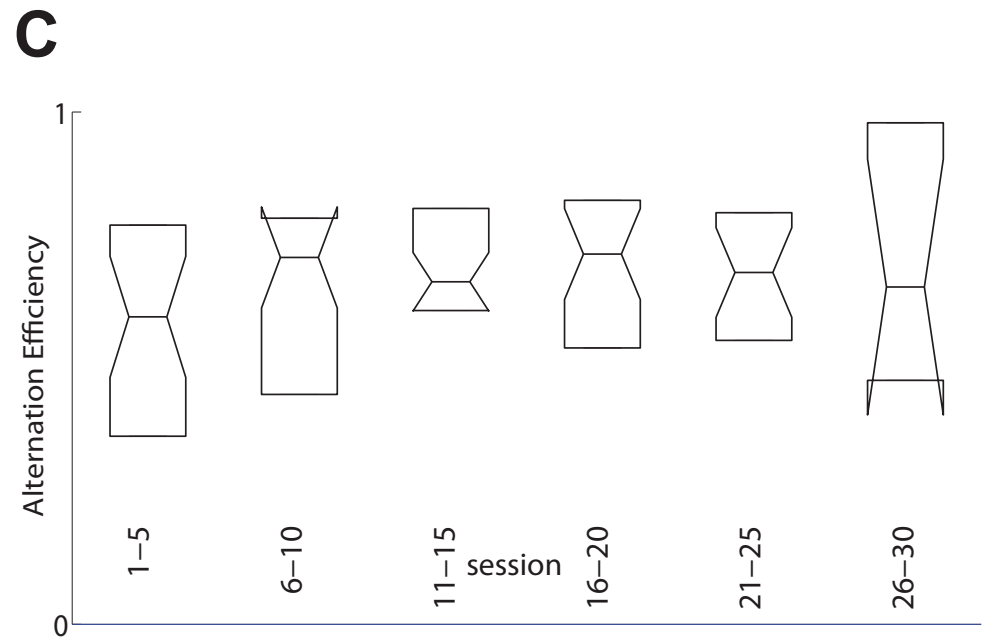
Supplemental Figure 2

Associated with Figure 3.

A: The rate of SWR events increased over sessions. This increase occured during all three phases, but was most consistent during the exploitation phase. The three inset panels show the same axes as the main panel, but only for laps restricted to each phase (investigation, titration, exploitation). **B:** The amount of VTE (measured by zIdphi) decreased over sessions. This decrease was driven entirely by changes during the exploitation phase. The three inset panels show the same axes as the main panel, but only for laps restricted to each phase (investigation, titration, exploitation). **C:** The alternation efficiency (see **supplemental methods**) increased over sessions. Boxplots show IQR (box), median (line), and standard error of the median (notch).


A
Overall
2
0
#Swr/s
1–5
6–10
11–15
16–20
21–25
26–30
session
Investigation
Titration
Exploitation

B
2
−2
zIdPhi
Overall
1–5
6–10
11–15
16–20
21–25
26–30
session
Investigation
Titration
Exploitation

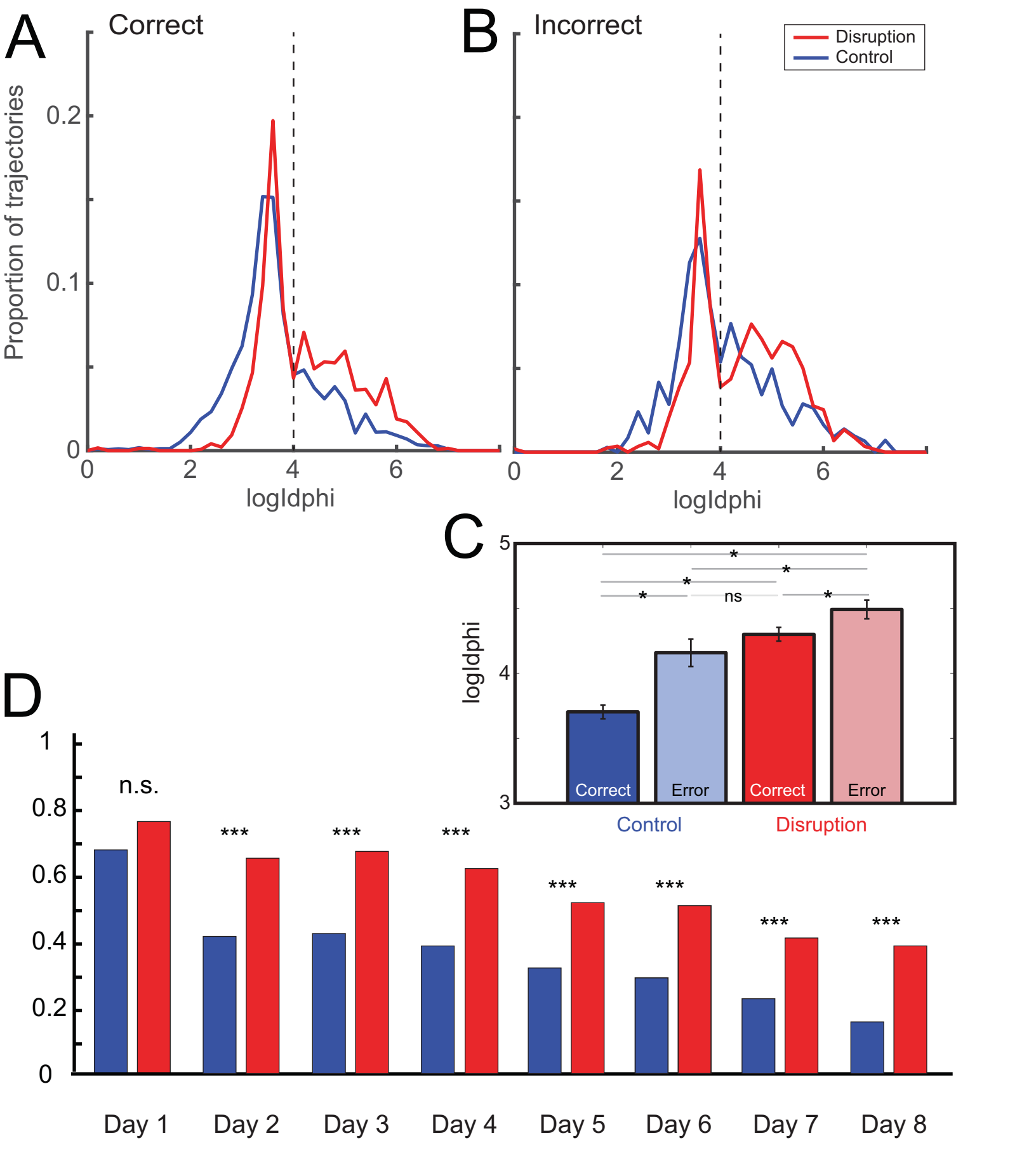

C
1
0
Alternation Efficiency
1–5
6–10
11–15
16–20
21–25
26–30
session

Supplemental Figure 3

Associated with Figure 4.

Effect of SWR disruption on VTE is not due to increase in errors, and persists across learning and performance of the alternation task. A. To determine whether the increase in VTE associated with SWR disruption was simply the result of more outbound errors, we separated our analysis of the effect of SWR disruption on VTE events into correct and incorrect trials. Distribution of lnIdphi limited to correct trials still showed an increase in high lnIdphi choice point passes for disruption animals as compared to control ($p < 10^{-6}$). **B**. We also saw an increase in high lnIdphi choice point passes for disruption animals during incorrect trials ($p < 10^{-10}$). **C**. There was a significant interaction ($p < 0.001$, see text) between the control / distruption groups and correct / error trials. Error bars show SEM. Asterisks indicate significant comparisons $p < 0.05$ as measured by a post-hoc Tukey test.

D. SWR disruption was maintained throughout the experimental period, which allowed us to ask what impact SWR disruption had on VTE behavior over learning. We saw a general decrease in VTE behavior over learning for both control and disruption animals, in agreement with delay discounting task. On the first day of task acquisition (when animals explored the novel W-track), the SWR disruption group matches control in proportion of VTE events. A difference in proportion of VTE events in SWR disruption animals was seen on all days except the first day. Differences were seen even on the final days (days 7 and 8) when animals performed mostly above chance-level (Jadhav et al., 2012).

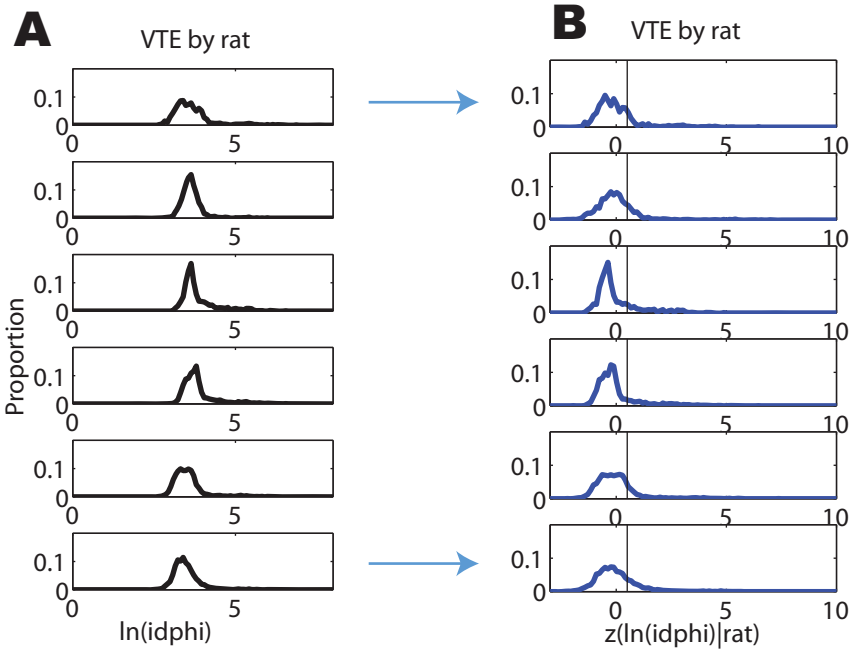

A
Correct
B
Incorrect
Disruption
Control
Proportion of trajectories
0.2
0.1
0
0
2
4
6
logIdphi
C
5
4
3
logIdphi
*
*
*
*
*
ns
Correct
Error
Correct
Error
Control
Disruption
D
1
0.8
0.6
0.4
0.2
0
n.s.
***
Day 1
Day 2
Day 3
Day 4
Day 5
Day 6
Day 7
Day 8

Supplemental Figure 4

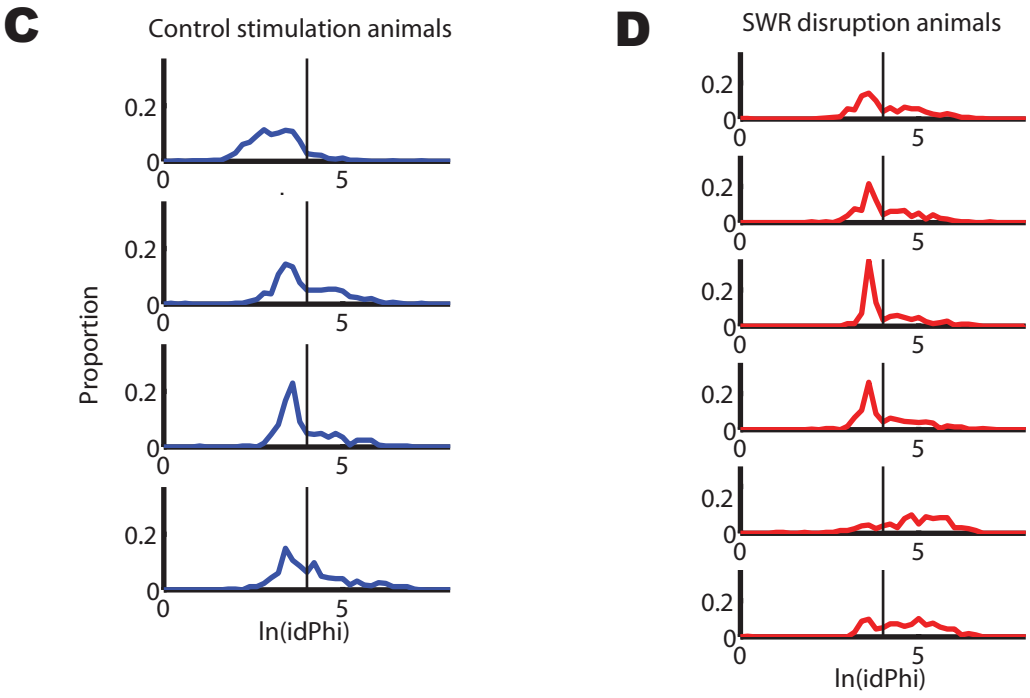
Associated with Figures 1 and 4.

Normalization procedures applied to the two experiments. The questions associated with the DD task require comparisons between laps occuring in each animal, so we normalized the lnIdphi measures (**A**) by z-scoring each measurement within rat (**B**). Because the questions associated with the W-maze task require between-animal comparisons, we did not normalize within-rat (which can obscure differences) and used the lnIdphi measure directly. See **supplemental methods**.

Experiment 1: VTE and SWR interactions on the DD task.
(within rat comparisons require within-rat normalization)



Experiment 2: Effect of SWR disruption on VTE on the W-maze task
(between rat comparisons preclude within-rat normalization)



Supplemental Material (5): Table 1

Statistics for changes observed over days in the DD task. We applied ANOVAs to each analysis, taking into account phase of the task and session number, and measuring their interaction. For each phase, for each process, we measured the slope using a single-variable linear regression and measured the probability that the slope as measured was different from zero.

VTE, change over sessions

ANOVA: effect of phase $p=10^{-27}$, effect of session $p=10^{-76}$, interaction $p=10^{-29}$

Phase	Slope	p-value (>0)
OVERALL	-0.19	$p<10^{-93}$
Investigation	-0.13	$p<10^{-05}$
Titration	-0.09	$p<0.001$
Exploitation	-0.21	$p<10^{-82}$

SWR/s, change over sessions

ANOVA: effect of phase $p=10^{-20}$, effect of session $p=10^{-75}$, interaction $p=10^{-100}$

Phase	Slope	p-value (>0)
OVERALL	+0.43	$p<10^{-100}$
Investigation	+0.46	$p<10^{-66}$
Titration	+0.39	$p<10^{-56}$
Exploitation	+0.42	$p<10^{-100}$

Alternation efficiency, change over sessions

ANOVA: effect of session $p<0.025$

Phase	Slope	p-value (>0)
OVERALL	+0.16	$p<0.05$

Supplemental Material: Experimental Procedures

DD task

Subjects. Six adult male Fisher 344 Brown Norway rats (Harlan, Indianapolis, IN) age 8-12 months at the beginning of the experiment were trained on the Delay Discounting (DD) task. Rats were individually housed on a 12 hour light / dark cycle and food restricted to no more than 80% of ad lib weight with water available ad lib. All experimental procedures were conducted in compliance with National Institute of Health guidelines for animal experimentation and approved by the Institutional Animal Care and Use Committee at the University of Minnesota.

Behavioral Task and Training The behavioral training, task design, and lap categorization have been described previously (Papale et al., 2012). At the same time each day (±2hr), rats were trained to run 100 laps / session on a T-maze version of an adjusting-delay discounting task. Each session, one reward site would deliver a larger-later reward (3×45mg sucrose pellets, Test Diet) and the other would deliver a smaller-sooner reward (1×45mg pellet). Rats were required to wait from 1-30s prior to delivery of the larger-later reward and 1s prior to delivery of the smaller-sooner reward. The delay preceding the larger-later reward was adjusted based on the decisions of the rat. Repeated choices to the larger-later side increased the delay by 1s on the subsequent lap, while repeated choices to the smaller-sooner side decreased it by 1s on the subsequent lap. These are called *adjustment* laps. *Alternation* between sides left the delay unchanged. Delays were accompanied by a tone-sequence countdown with each delay matched to a specific pitch. The tones began after leaving the choice point, so analysis of choice point behavior (and associated neurophysiology) did not include tone cues.

Rats ran one session per day. Each rat ran 30 sessions before being implanted with tetrodes for neural ensemble recordings, were then re-trained as tetrodes were lowered into the hippocampus. Once tetrodes were recording hippocampal neurons, rats ran an additional 30 sessions (1 session / day). The recording sessions are analyzed here.

Rats were tracked by LEDs on the recording headstage from an overhead camera with a resolution of 0.17 cm / pixel and at a speed of 60Hz.

Vicarious Trial and Error (VTE) behavior was quantified with the zIdphi measure (the z-scored, integrated absolute angular velocity ($d\phi$), see Papale et al., 2012). In short, zIdphi measures the integrated angular velocity of the orientation of motion. After calculating the Idphi measure, we took the $\ln$ of it to reduce the skewness. Because the first experiment (DD) depended on comparisons between laps occuring in each animal, we then z-scored the $\ln$ Idphi measure using means and standard deviation calculated from all laps for a given rat. These transformations are shown in **Fig. S4**. In order to identify a threshold to separate VTE laps from non-VTE laps, we found the 5 most populous bins in the histogram shown in **Fig. 1D**, took the mean zIdphi of those 5 bins, then took all the samples below that mean, reflected them around the mean, and used that pseudo-sample to define an expected Gaussian distribution of non-VTE laps. The observed distribution diverged from the expected Gaussian at zIdphi =0.5. Therefore, for the DD task, VTE laps were classified as those with zIdphi > 0.5.

The Idphi measure includes both speed through the choice point as well as variability in the path through the choice point. While Idphi does not provide a sharp decision boundary between VTE and not, previous studies have found that choice point passes with Idphi above the decision point tend to be VTE laps with pauses and a high behavioral variability, while those below tend to be fast and regular passes with a consistent speed indicative of more ballistic movements (van der Meer and Redish, 2010; Blumenthal et al., 2011; Steiner and Redish, 2012; Stott and Redish, 2014; Amemiya and Redish, 2016). This decision variable (defined a priori from previous studies) was sufficient to identify significant differences in the hippocampal ensemble representation (see **Fig. 2.**)

Behavioral analysis: Rats on the DD task typically show behaviors indicative of three phases: an investigation phase, a titration phase, and an exploitation phase (Papale et al., 2012). Importantly, these phases were not imposed on the task, but rather measured from the rat's behavior. *Investigation* was defined as laps before the first adjustment lap. *Exploitation* was defined as laps following the first time the rat reached the indifference point for that day. The **indifference point** was measured as the mean adjusted delay on the delayed side over the final 20 laps of the session. *Titration* was defined as the laps between investigation and exploitation.

Alternation efficiency was defined as the number of alternation laps when the delay was within 3s of the final measured indifference point divided by the total number of alternation laps. **Path stereotypy** was measured by first defining laps as proceeding from

departure from one feeder (measured as the time crossing a spatial threshold) and arrival at the next (measured, again, as the time crossing a spatial threshold). The 2D $\langle x, y \rangle$ path was resampled to 1000 samples. For each pair of laps (L_i, L_j), the distance between the paths was defined as the mean distance between the matched samples. Thus paths that are both spatially and temporally matched will have small distance measures. Separate analyses were done for left-left (LL), right-right (RR), left-right (LR), and right-left (RL) laps. Analyses were done for each session, and then averaged across sessions. Lap pairs from different paths (LL x RR, LR x RL, etc) were defined as `NaN` and did not enter into the analyses.

Neural Recording. Rats were implanted with 12-tetrode hyperdrives (Kopf) using standard procedures targeting right dorsal CA1 (-3.8mm AP, -3.0mm ML). Over the next 10-14 days, tetrodes were lowered at a rate of 20-320μm / day, with larger adjustments initially and finer adjustments as target depth was approached (1200-1800μm). Local field potentials were used to align tetrodes as they approached CA1, and CA1 depth was confirmed when putative pyramidal cells that spiked during flat sharp wave ripples were observed.

Action potentials were recorded at a rate of 32kHz within a 1ms window if, on any of the four channels, voltage exceeded a user-defined threshold. Spikes were then filtered (600-9000Hz) and digitized (Neuralynx, Cheetah, Tuscon, AZ). The local field potential voltage was sampled from one channel per tetrode at 2kHz and bandpass filtered (1-475Hz). Spike sorting was carried out offline using an automatic k-means clustering algorithm (KlustaKwik, Harris et al.), and then manually using waveform parameters in a multi-dimensional space (MClust 3.5, Redish et al.).

SWR Analysis. SWRs were analyzed while rats paused at the feeders before beginning the next lap. Pause time was taken from the time of reward delivery to the time that rats left the feeder zone, defined as a circle of pixels recorded from the overhead camera. SWR rate was defined as the number of SWR at the feeder site normalized by the pause time.

SWRs were detected using the tetrode with the most cells per session. The local field potential from that tetrode was bandpass filtered from 150-250Hz and then a Hilbert transform was applied. The instantaneous amplitude obtained from the Hilbert was

z-scored and global extrema were computed with those greater than 2.5 standard deviations above mean amplitude were considered SWR. A SWR that occurred less than 250ms from a prior event was discarded.

Bayesian Decoding. A Bayesian algorithm was applied to analyze the information content of CA1 ensembles during SWR and choice point passes (Zhang et al., 1998). If a cell had an average firing rate of less than 10Hz, its firing rate was binned into 50ms windows. Tuning curves were defined from the cell's occupancy-normalized firing in a 56×56 spatial grid using data from the training set. All spikes during theta epochs and those not occurring during the ± 125ms window around SWR events were used for the training set. The posterior probability was then computed for each 50ms time bin, producing a probability distribution across the 56 x 56 grid at each time step. Analysis zones were defined as four rectangles (see **Fig. 2B**). Four of the six rats had sufficiently large neural ensembles (> 10 cells / session) to do Bayesian decoding analyses. All six rats were included in the SWR and behavioral analyses.

Spectral Analyses. Spectral analyses were done using standard spectral methods (spectrogram function, Matlab) with a Hamming window sized to be able to access 2 Hz. Power spectral density plots were generated by averaging across these spectra to remove the temporal dimension. Frequency-frequency autocoherence plots (Masimore et al., 2004) were generated by correlating across these spectra to remove the temporal dimension.

W-Maze task

Animals. 14 male Long Evans rats weighing 450-600 grams were used in this study. All procedures were approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco and conformed to National Institutes of Health guidelines.

Behavior. Animals were food deprived to 85-90% of their baseline weight and were pre-trained to alternate between two ends of a linear track for liquid food reward (evaporated milk) dispensed automatically. Animals were tested on the novel W track for

8 days, two 15min sessions per day. Food wells at the end of each arm automatically dispensed liquid food reward according to alternation rules previously described (Jadhav et al., 2012). In brief, center arm visits were rewarded when the previously rewarded (n-1) food well was located in one of the side arms. A side arm well visit was rewarded when the previously rewarded food well (n-1) was located in the center arm, and the previously visited side arm (n-2) was the opposite side arm, that is, the animal had successfully alternated visiting side arms. Repeated visits to the same reward well and incorrect alternations were not rewarded. At the conclusion of each run session, animals were placed in a familiar rest box for 15min.

Stimulation properties. Surgical procedures to implant electrodes were as previously described (Jadhav et al., 2012). In addition to 12-tetrodes targeting CA1 (-3.6 mm AP,+2.2 mm ML), one or two tungsten stimulation electrodes were implanted targeting vHC (-1.3 mm AP, $\pm$1 mm ML) ipsilaterally or bilaterally to CA1 recording tetrodes. Stimulation pulses were biphasic and 0.2 ms in duration, with amplitude calibrated to induce 100 ms of inhibition of CA1 multiunit activity (ranging from 40-180μA).

Real time SWR detection and disruption. For full details, see (Jadhav et al., 2012). Briefly, we chose local field potentials for 5-6 tetrodes, filtering in the ripple band (20 tap band-pass IIR filter, 100-400 Hz). To establish a threshold for disruption, we calculated smoothed means and standard deviations of the absolute value of the LFP on each tetrode. The threshold for disruption was then set to 4-6 s.d. above the mean on at least two tetrodes. For control animals, we introduced a latency of 150-200 ms between onset of stimulation and SWR detection, effectively decoupling the two and maintaining SWR content.

Position tracking and reconstruction. Illuminated IR LEDs were attached to the front and back of the animals' headstage during recording for position and speed reconstruction. Rat behavior sessions were then tracked via fixed overhead monochrome CCD camera at 30 Hz with a 0.45 cm / pixel resolution. We reconstructed position using a semi-automated thresholding analysis using custom software (MATLAB, Mathworks). Position was then smoothed using a non-linear method (Jadhav et al., 2012) and used to separate the behavior into individual trials (behavior trajectories) based on well position

and W alternation sequence order (returning to center well, inbound; proceeding to alternating side arm, outbound). Trajectories were classified as correct or incorrect based on correspondence between alternation sequence order and reward wells visited. Outbound, memory dependent trajectories were selected for further analysis.

Choice zone identification. The choice zone of each trajectory was defined as a circle with a radius of 15 cm, centered at the intersection point of the middle maze arm and the remainder of the W track. The first instance of the animal crossing into the choice zone and the first instance of the animal leaving the choice zone were considered the start and end of the choice zone trajectory. These positions were used for subsequent ldphi and VTE calculations.

lnIdphi and VTE calculation. We calculated ldphi for each choice-zone trajectory, using equivalent methods to that used in the DD task. Because the W-maze analyses depend on between-animal comparisons, we did not normalize the VTE measures within rats, but used the lnIdphi measure directly. For the W-maze task, VTE laps were classified as those with a lnIdphi measure greater than 4.

lnIdphi and VTE behavior was quantified similarly for SWR disruption (n = 6) and control stimulation (n = 4) animals.