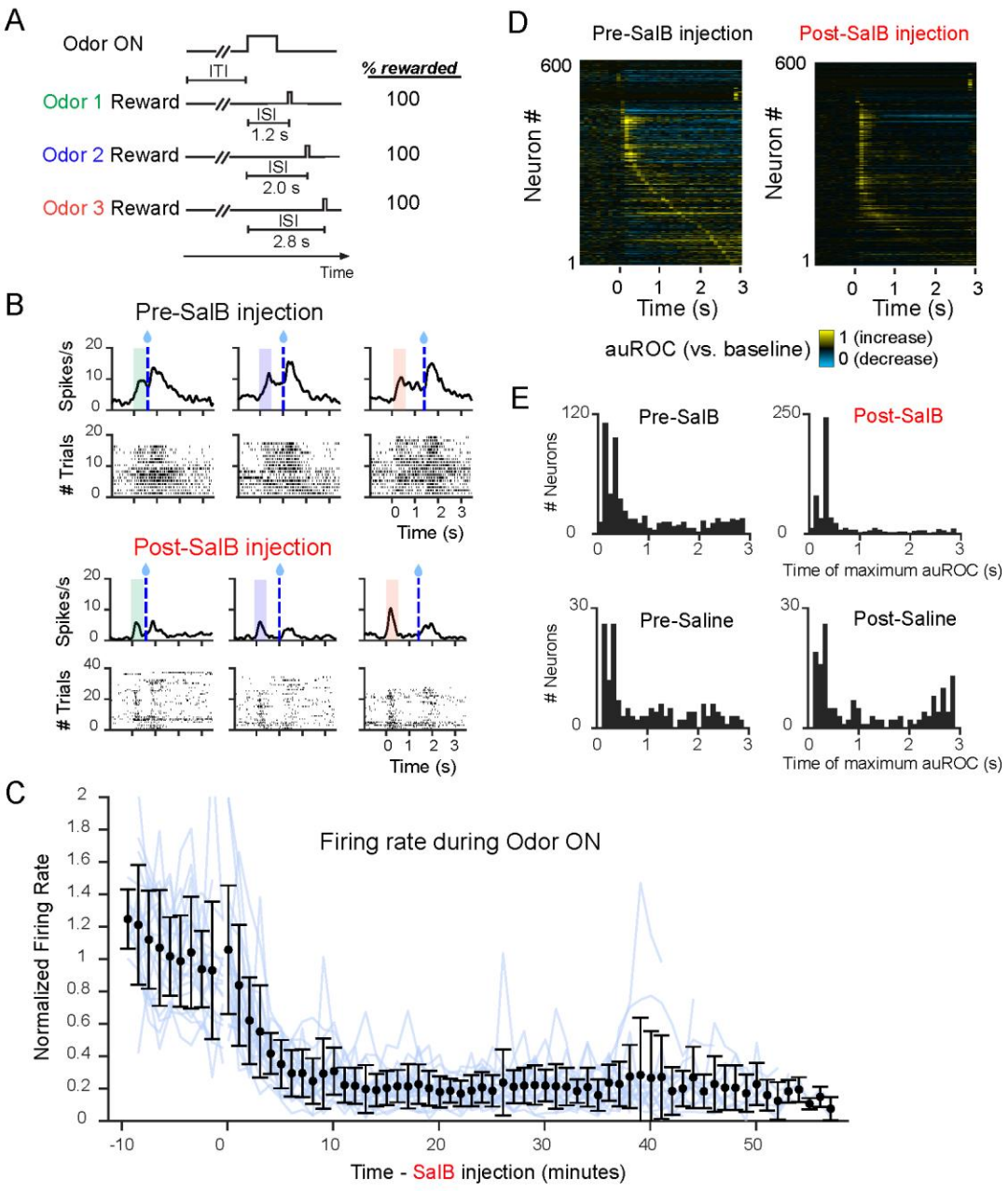

Figure S1. [KORD effect on Task-related activity], related to Figure 2

(A) The two mPFC-implanted animals were trained on Task 1 (Figure 1), plus an Odor that had a constant ISI of 2.0s (Odor 2). Data shown in this Figure are displayed only for constant ISI odors (1-3) due to insufficient number of pre-injection trials for the other trial types. **(B)** Responses of a single example neuron before and after SalB injection. This neuron displayed sustained activity during the ISI, prior to SalB injection. Post-SalB injection, the sustained activity (as well as baseline firing rate) decreased. **(C)** Firing rate during odor duration (0-1s following odor onset) for all inhibited neurons in the SalB condition from **Figure 2F**. Although some task-related activity was spared following SalB injection, as shown for the example neuron in **(B)**, this task-related activity also tended to decrease following SalB injection. **(D)** Area under the ROC curve (auROC) computed for Odor 3 responses (ISI = 2.8s) for each mPFC neuron, before and after SalB, arranged in order from earliest to latest maximum response. Visibly fewer neurons show a profile of peak activation later in the time interval after SalB injection. **(E)** Histogram of all neurons' time bin of maximal activity, computed by auROC. Wilcoxon rank-sum test revealed the distribution was significantly shifted post-injection in the SalB condition ($z = 3.93$, $p = 8.6 \times 10^{-5}$) but not the Saline condition ($z = 1.72$, $p = 0.084$).

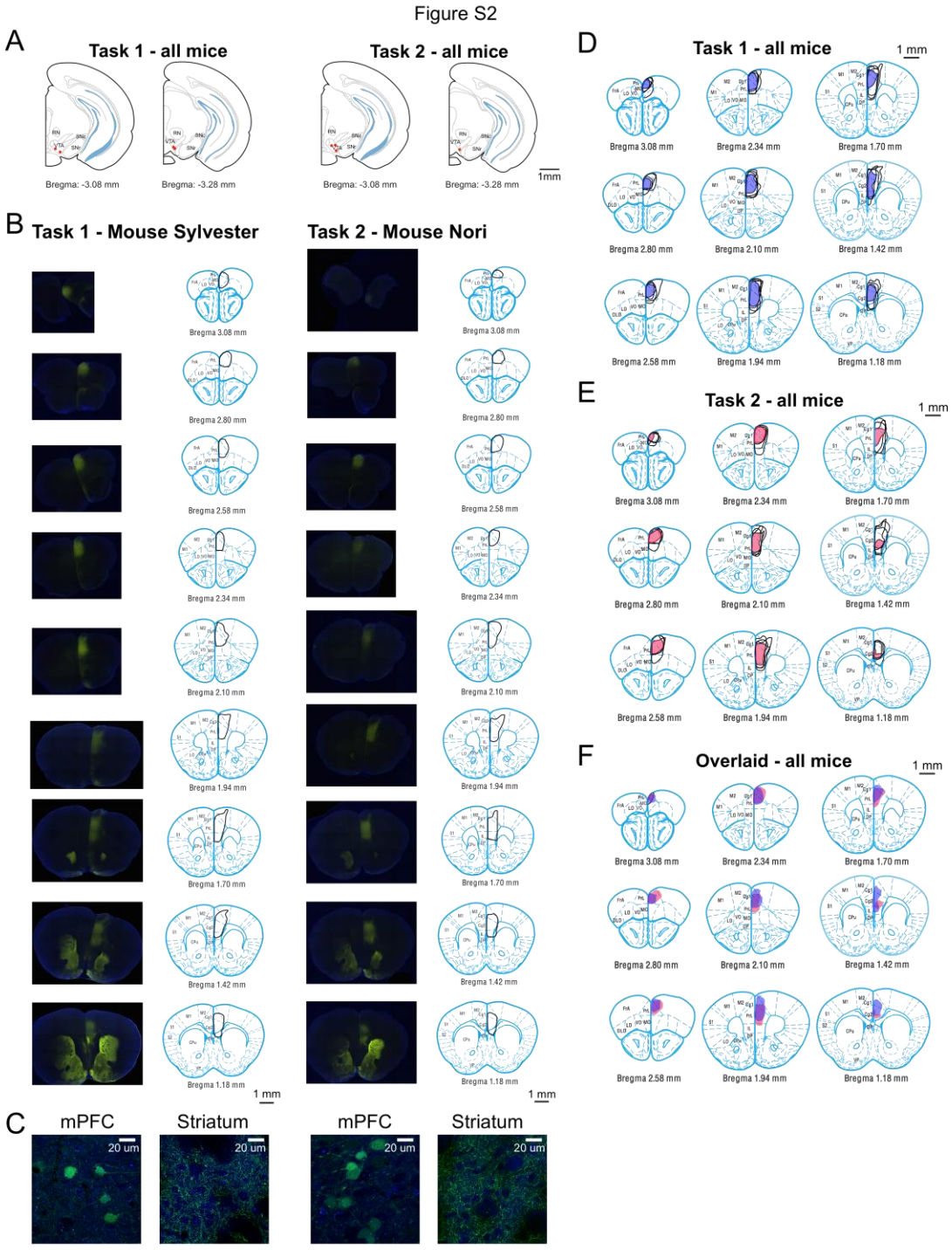

Figure S2. [VTA recording sites and KORD expression], related to Figure 3

(A) Recording sites in Task 1 and Task 2 animals. **(B)** Raw histology images for two representative mice from Tasks 1 and 2, respectively. **(C)** While striatal labeling appears bright in (**B**), 63x confocal images reveal this striatal labeling to consist predominantly of axons, in contrast to cell bodies that are labeled in the mPFC. **(D-E)** Areas outlined in black denote the site of KORD expression in individual Task 1 and Task 2 animals. Blue areas in (**D**) indicate minimum area covered in all Task 1 animals, and red areas in (**E**) indicate minimum area covered in all Task 2 animals. **(F)** Minimum areas covered in Task 1 and 2 animals, overlaid.

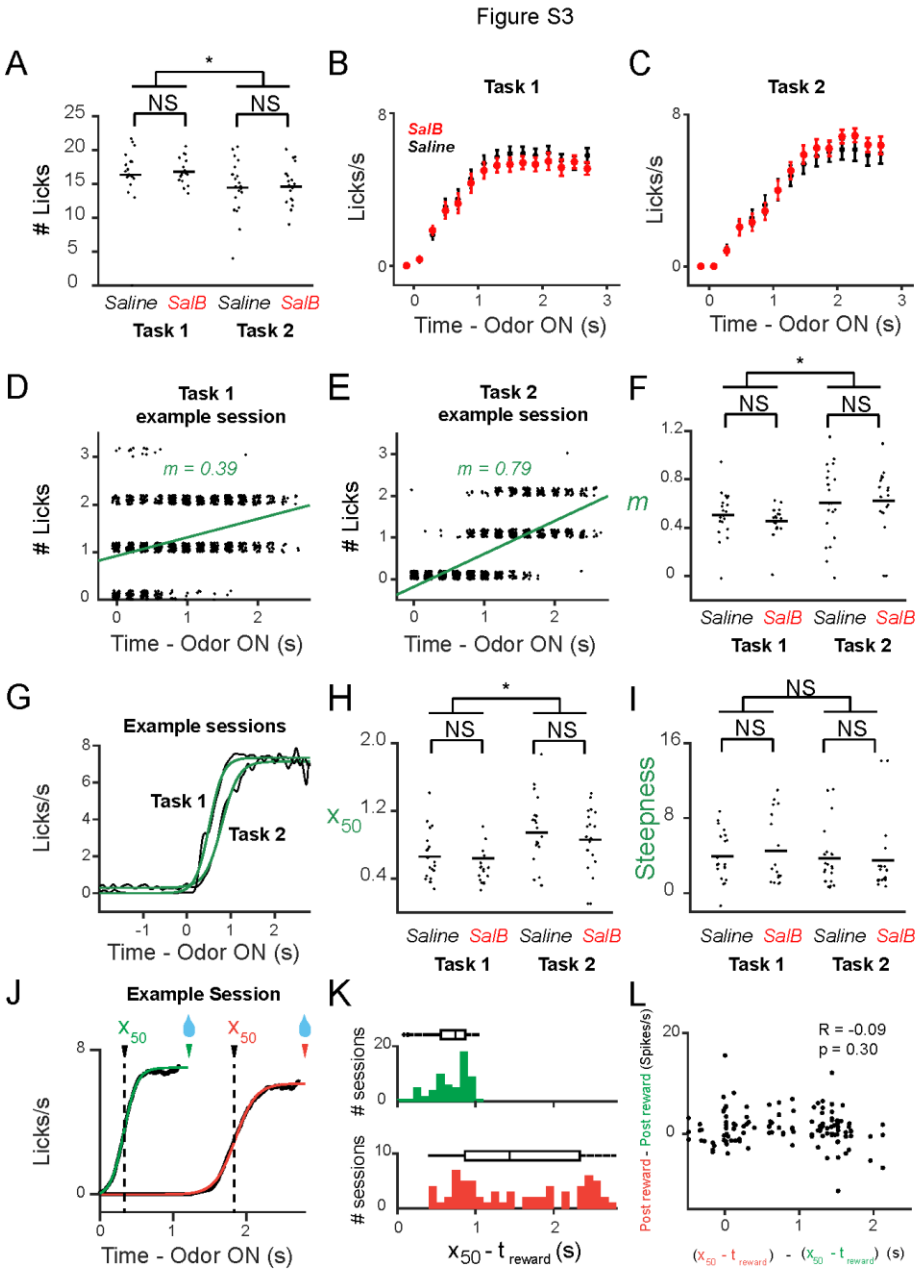

Figure S3. [Analysis of animal behavior in Tasks 1 and 2, Saline and SalB conditions], related to Figure 3

(**A**) Comparing the number of licks from odor onset until reward was received (only for trials in which reward arrived at 2.8s, to capture licking throughout the entire possible duration of the ISI). The number of licks were significantly different between Task 1 and Task 2 (# anticipatory licks for Odor A in Task 1 ≠ Task 2; $F_{1,76} = 6.7$, $p = 1.16 \times 10^{-2}$, one-way analysis of variance (ANOVA)), but not between Saline and SalB conditions, within each task (# anticipatory licks for Odor A in Task 1 SalB ≠ Task 1 Saline; $F_{1,36} = 0.16$, $p = 0.69$; # anticipatory licks for Odor A in Task 2 SalB ≠ Task 2 Saline; $F_{1,38} = 0.01$, $p = 0.91$, one-way analysis of variance (ANOVA)). (**B,C**) Comparing the lick rate after Odor A onset until reward is received, in 200ms epochs, between Saline and SalB conditions. We performed 1-way ANOVAs comparing the lick rate between the Saline and SalB condition for each epoch (14 total ANOVAs), for each task. Lick rates were not significantly different in any epoch between Saline and SalB conditions (lick rate for any given epoch after Odor A in Task 1 Saline condition ≠ Task 1 SalB condition; $F_{1,36} < 4.0$, $p > 0.40$ for all comparisons; lick rate for any given epoch after Odor A in Task 2 Saline condition ≠ Task 2 SalB condition; $F_{1,38} < 4.0$, $p > 0.40$ for all comparisons, one-way analysis of variance (ANOVA) – no correction for multiple comparisons). (**D,E**) Best-fit lines relating the number of licks to time following Odor A (but before reward onset) were plotted for behavior during individual recording sessions in Tasks 1 and 2. The best-fit lines and slopes (m) for these two example sessions are shown in green. (**F**) The slopes were

Figure S3 (*continued from previous page*) significantly different between Task 1 and Task 2 (slopes in Task 1 ≠ Task 2; $F_{1,76} = 6.1$, $p = 1.6 \times 10^{-2}$, one-way analysis of variance (ANOVA)), but not between Saline and SalB conditions, within each task (# anticipatory licks for Odor A in Task 1 SalB ≠ Task 1 Saline; $F_{1,36} = 0.87$, $p = 0.36$; # anticipatory licks for Odor A in Task 2 SalB ≠ Task 2 Saline; $F_{1,38} = 0.04$, $p = 0.84$, one-way analysis of variance (ANOVA)). (**G**) Best-fit sigmoid curves relating the lick rate to time following Odor A (but before reward onset) were plotted for behavior during individual recording sessions in Tasks 1 and 2. The best-fit curves for these two example sessions are shown in green. (**H**) The timepoints of halfway-to-maximum lick rates were significantly different between Task 1 and Task 2 (x_{50} in Task 1 ≠ Task 2; $F_{1,76} = 7.69$, $p = 7.0 \times 10^{-3}$, one-way analysis of variance (ANOVA)), but not between Saline and SalB conditions, within each task (x_{50} in Task 1 SalB ≠ Task 1 Saline; $F_{1,36} = 0.02$, $p = 0.90$; x_{50} in Task 2 SalB ≠ Task 2 Saline; $F_{1,38} = 0.42$, $p = 0.52$, one-way analysis of variance (ANOVA)). (**I**) The steepness of the sigmoid fits were not significantly different between Task 1 and Task 2 (steepness in Task 1 ≠ Task 2; $F_{1,76} = 0.59$, $p = 0.44$, one-way analysis of variance (ANOVA)), nor were they different between Saline and SalB conditions, within each task (steepness in Task 1 SalB ≠ Task 1 Saline; $F_{1,36} = 0.3$, $p = 0.59$; steepness in Task 2 SalB ≠ Task 2 Saline; $F_{1,38} = 0.05$, $p = 0.82$, one-way analysis of variance (ANOVA)). (**J**) Examples of the sigmoid fits for Odor B and C trials, from one example session. The Odor C licking trace is in orange, and the Odor B licking trace is in green. (**K**) The time of reward (t_{reward}) minus the time to half the maximum lick rate (x_{50}), for every individual session. We used this as a measure of temporal imprecision: the smaller $x_{50} - t_{reward}$, the more precisely the animal licked exactly when water was delivered. (**L**) Based on the behavioral imprecision computed in (**K**), we attempted to predict the increase in dopamine RPEs from Odor B to Odor C rewards. However, this analysis did not reveal a significant correlation. For this reason, we chose the Weber fraction in our model based on other animal timing studies, rather than behavior from our tasks.

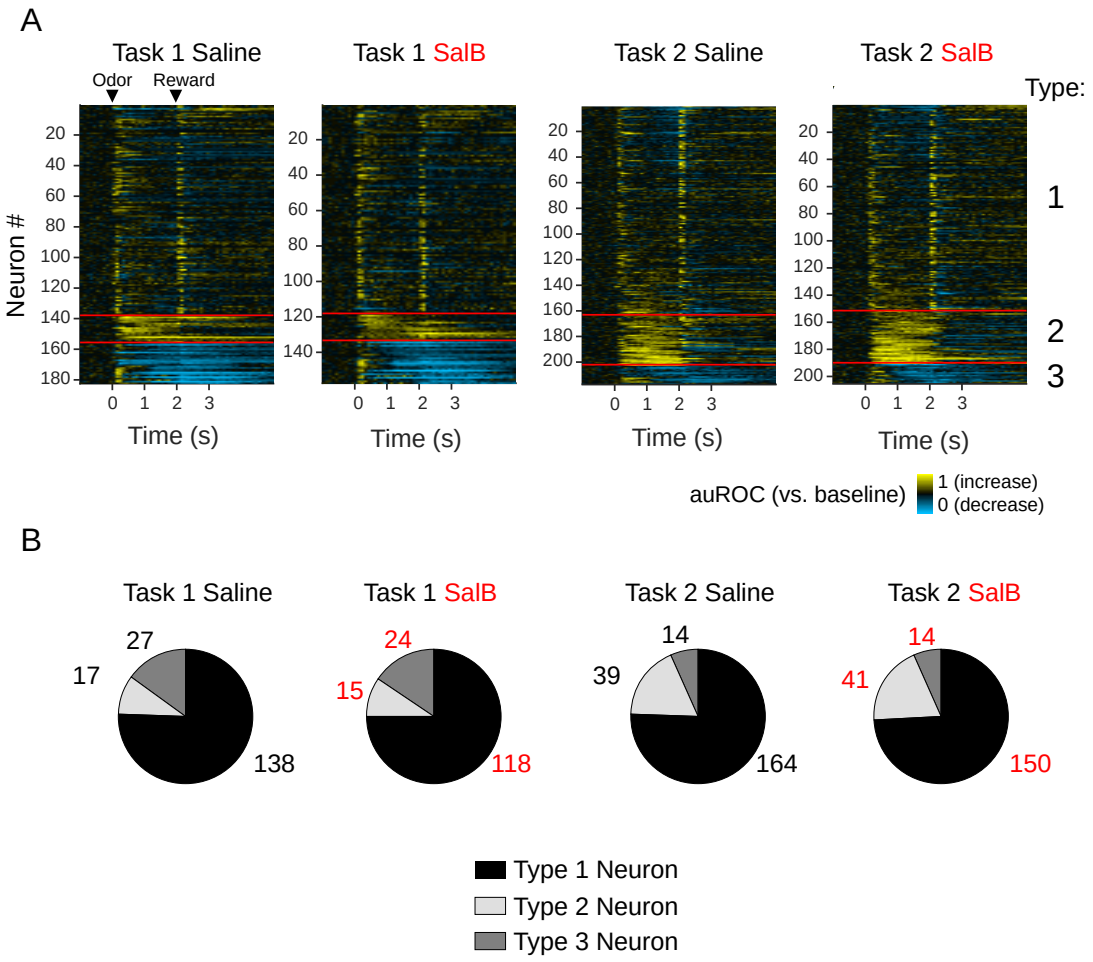

Figure S4. [Neurons recruited to the task in Tasks 1 and 2, Saline and SalB groups], related to Figure 3

(A) Area under the ROC curve (auROC) computed for Odor A responses (ISI = 2.0s) for each recorded VTA neuron (within 500um of an optogenetically-identified dopamine neuron), clustered into 3 groups by k-means clustering. Type 1 neurons show phasic responses to both cue and reward. Type 2 neurons show sustained positive responses spanning the time between cue and reward. Type 3 neurons show sustained negative responses spanning the time between cue and reward. (**B**) We did not find appreciable differences in the number of particular types of neurons recruited to the recording sessions, between Saline and SalB conditions.

Figure S5

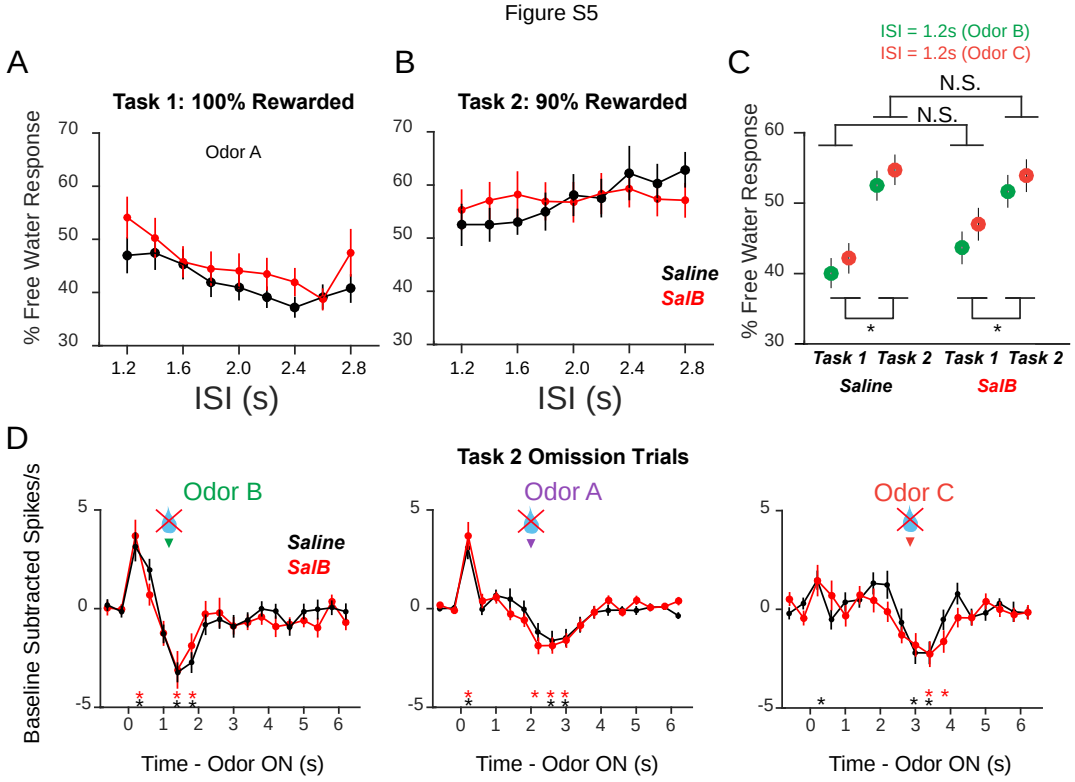

Figure S5. [Normalized RPEs], related to Figure 4

(A,B) The number of post-reward spikes fired over baseline following various Odor A deliveries was divided by the number of post-reward spikes fired over baseline following free water delivery, for Task 1 (**A**) and Task 2 (**B**). Saline data is in black, and SalB data is in red. (C) The number of post-reward spikes fired following Odor B and C deliveries was divided by the number of post-reward spikes fired following free water delivery, in both tasks and conditions. The magnitude of Odor B and C post-reward responses was not significantly different between Saline and SalB conditions (Normalized post-reward responses in Task 1 SalB ≠ Saline, $F_{1,162} = 2.4$, p = 0.12; Task 2 SalB ≠ Saline, $F_{1,168} = 0.06$, p = 0.81, one-way analysis of variance (ANOVA)). In both Saline and SalB conditions, the magnitude of Odor B and C post-reward responses significantly increased from Task 1 to Task 2 (Normalized post-reward responses in Saline Task 1 ≠ Task 2, $F_{1,158} = 19$, $p = 2.6 \times 10^{-5}$; SalB Task 1 ≠ Task 2, $F_{1,172} = 5.3$, p = 0.023, one-way analysis of variance (ANOVA)). (**D**) Baseline-subtracted reward omission responses, plotted in 200ms epochs. Asterisks indicate baseline-subtracted firing rates that significantly ($P < 0.05$) deviated from zero (Wilcoxon sign-rank test, Bonferroni corrected for multiple comparisons). Firing rates at each epoch, in each trial type, were not significantly different between Saline and SalB conditions ($P > 0.05$ for all comparisons, Bonferroni corrected for multiple comparisons).

Figure S6

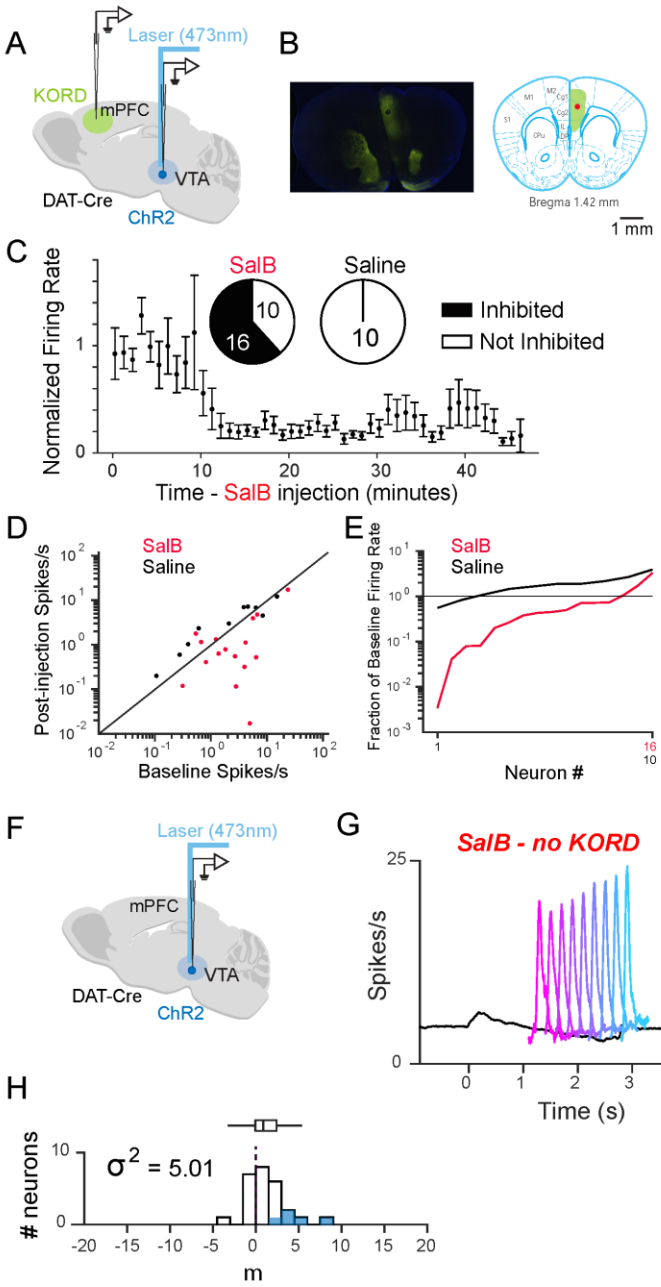

Figure S6. [Additional KORD controls], related to Figure 4

(A) To verify that KORD was efficacious in Task 1, we recorded from the mPFC of one animal trained and recorded on Task 1. Schematic showing 1 tetrode implanted in KORD-expressing mPFC, and the other tetrodes implanted in the VTA for dopamine recording. **(B)** Coronal section from the animal schematized in (**A**), showing KORD expression in green. The hole in the tissue is the electrolytic lesion created at the end of the tetrode, following completion of the experiment, showing that the tetrode was located in a region that expressed KORD. **(C)** Normalized firing rates after injecting the animal from (**B**), for all inhibited mPFC neurons (suppressed firing rate to less than half of baseline firing rate – here, baseline was taken to mean the 5 minutes post-SalB injection). Each neuron's firing rate was normalized to its baseline firing rate. **(D)** Baseline firing rate versus >15 minutes post-injection firing rate, on a log-log scale, for the animal in (**B**). **(E)** >15 minutes post-injection firing rate as a proportion of baseline firing rate, for each neuron, for the animal in (**B**). Neurons are rank-ordered from most-suppressed to most-excited. (**F**) To confirm that SalB alone did not affect dopamine RPEs, we recorded dopamine neurons in 2 animals trained on Task 2 that were injected with SalB, but lacking KORD. **(G)** Average non-normalized PSTH for all 26 dopamine neurons recorded during Odor A trials in Task 1, Saline condition. Post-reward firing: $F_{8,200} = 2.18$, p = 0.03, 2-way ANOVA; factors: ISI, neuron; Pre-reward firing: $F_{8,200} = 14.98$, $p = 1.1 \times 10^{-6}$. **(H)** 18/26 neurons displayed positive temporal modulation.

Figure S7

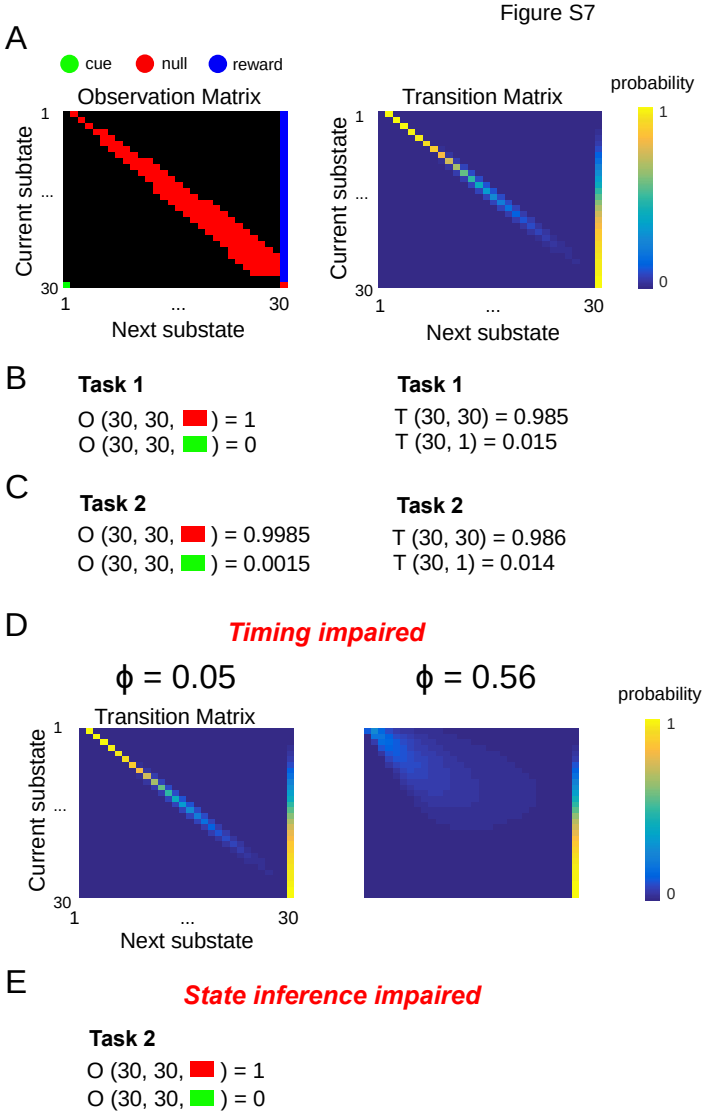

Figure S7. [Computational model details], related to Figure 7

(A) Observation and Transition matrices for belief state TD model (see *Methods*). The transition matrix was blurred by a Weber fraction of 0.05 (per sub-state; 0.25 per second). **(B,C)** The only differences in model values between Tasks 1 and 2 are noted. Most notably, the observation matrix value specifying the likelihood of observing a cue, given the transition back into the ITI, is 0 in Task 1 and non-zero in Task 2. **(D)** Impaired timing was simulated by blurring the transition matrices by a larger Weber fraction (see *Methods*). **(E)** Impaired hidden state inference was simulated by eliminating the non-zero likelihood of observing cue given an ITI→ITI transition, abolishing the model's ability to acknowledge the possibility of this hidden state transition during an omission trial.

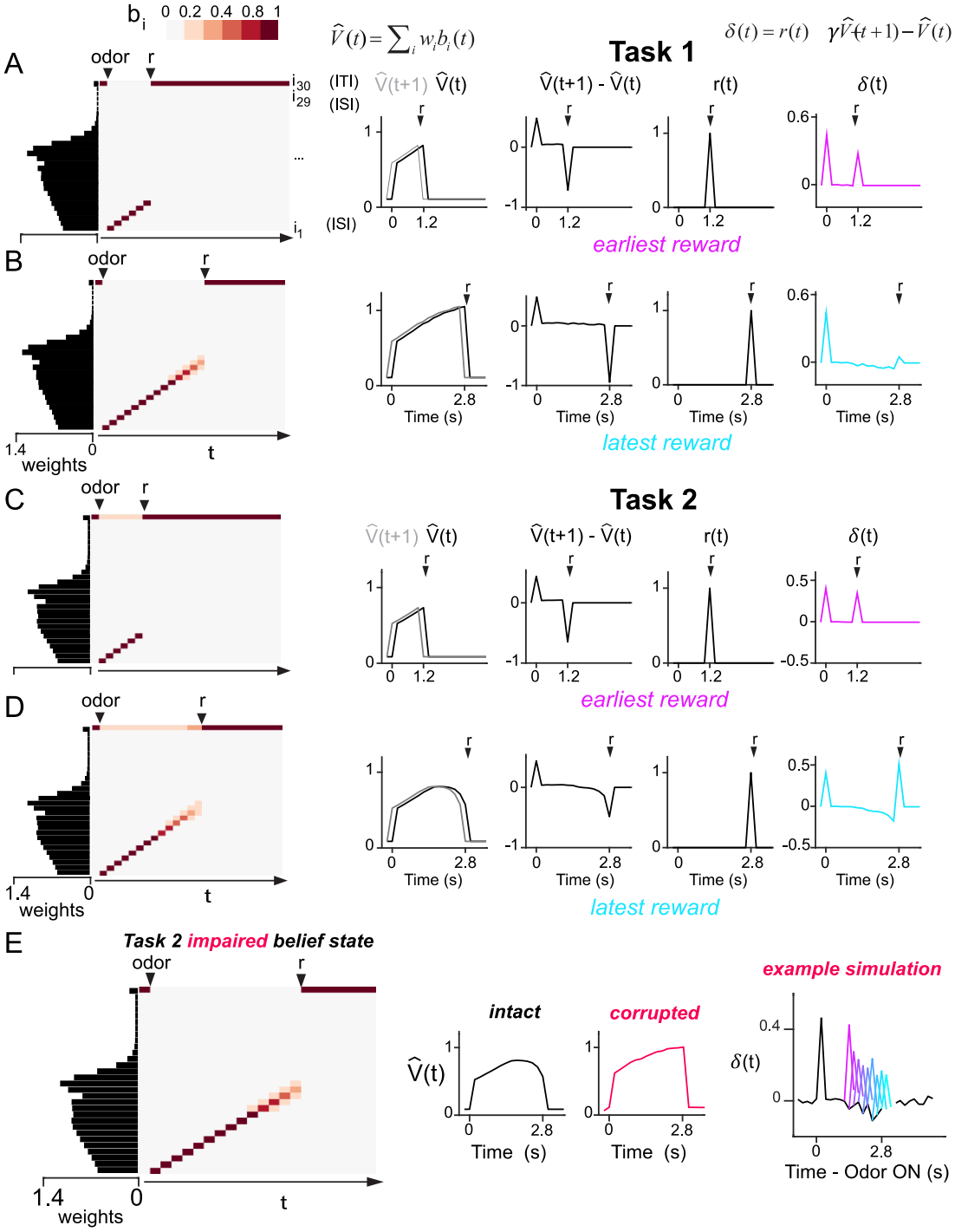

Figure S8. [RPE calculation in Belief State TD model], related to Figure 7

(A,B) As time elapses following odor onset in Task 1, the belief state proceeds through ISI sub-states with temporal estimation that becomes slightly less precise, but sums to 100% at each timestep. Later ISI sub-states accrue greater weights. Estimated value is approximated as the dot product of belief state and weight, producing a ramping value signal that increasingly suppresses $\delta(t)$ for longer ISIs. **(C,D)** As time elapses following odor onset in Task 2, the belief state comprises a probability distribution that gradually decreases for ISI sub-states and gradually increases for the ITI sub-state (i_{30}). This produces a value signal that declines for longer ISIs, resulting in the least suppression of $\delta(t)$ for the latest ISI. **(E)** The impaired Task 2 belief state no longer yields to the ITI as time elapses. The value signal in Task 2 simulations with a corrupted belief state ramps over time, similar to Task 2. Due to this ramping value signal, example simulations with corrupted belief states in Task 2 show negative temporal modulation over time.