Estrogen modulates reward prediction errors and reinforcement learning

Check for updates

Carla E. M. Golden [1], Audrey C. Martin[1], Daljit Kaur[1], Andrew Mah[1], Diana H. Levy[1], Takashi Yamaguchi [2], Amy W. Lasek [3], Dayu Lin [2], Chiye Aoki[1] & Christine M. Constantinople [1]✉

Gonadal hormones act throughout the brain and modulate psychiatric symptoms. Yet how hormones influence cognitive processes is unclear. Exogenous 17β-estradiol, the most potent estrogen, modulates dopamine in the nucleus accumbens core, which instantiates reward prediction errors (RPEs), the difference between received and expected reward. Here we show that following endogenous increases in 17β-estradiol, dopamine RPEs and behavioral sensitivity to previous rewards are enhanced, and nucleus accumbens core dopamine reuptake proteins are reduced. Rats adjusted how quickly they initiated trials in a task with varying reward states, balancing effort against expected rewards. Nucleus accumbens core dopamine reflected RPEs that influenced rats' initiation times. Higher 17β-estradiol predicted greater sensitivity to reward states and larger RPEs. Proteomics revealed reduced dopamine transporter expression following 17β-estradiol increases. Finally, knockdown of midbrain estrogen receptors suppressed sensitivity to reward states. Therefore, endogenous 17β-estradiol predicts dopamine reuptake and RPE signaling, and causally dictates the impact of previous rewards on behavior.

Estrogenic hormones, including 17β-estradiol, rhythmically fluctuate over the reproductive cycle in mammalian females and act on estrogen receptors, which can regulate protein expression via transcription and activation of signal transduction pathways[1–3]. These hormonal fluctuations support physiological imperatives such as homeostasis, fertility and sexual behavior[2], but, due to broad and diffuse mechanisms of action, also bind to receptors in brain regions that support cognitive behaviors, including decision-making.

Dopamine signaling in the nucleus accumbens core (NAcc) is modulated by application of exogenous 17β-estradiol in slices of ovariectomized animals[4,5]. The NAcc is a part of the ventral striatum that is innervated by midbrain dopaminergic neurons in the ventral tegmental area (VTA), and is required for learning cue–outcome associations[6] and determining the vigor of motivated behaviors[7]. Dopamine release in the NAcc is thought to instantiate reward prediction errors (RPEs) for reinforcement learning, which is a powerful theoretical framework for describing how animals learn through interacting with the environment[8].

In reinforcement learning, agents learn the value of taking actions in different states; those state–action values are iteratively updated based on RPEs, or the difference between received and expected rewards. A wealth of evidence suggests that dopamine release in the NAcc encodes RPEs[9–14]. Critically, manipulating dopamine release in this circuit has provided causal evidence that it acts as an RPE for reinforcement learning[15–17]. More precisely, converging synaptic inputs (for example, from cortex, hippocampus) onto medium spiny neurons, the principal cells of the striatum, are thought to convey the animal's state. Dopamine is thought to mediate plasticity at these synapses, such that animals will be more (or less) likely to take actions in a state that produced positive (or negative) RPEs. In actor–critic models of the basal ganglia, the NAcc is thought to act as a 'critic' that learns the expected future reward in a given state (that is, state value)[18,19], which is

[1]Center for Neural Science, New York University, New York, NY, USA. [2]Neuroscience Institute, New York University Grossman School of Medicine, New York, NY, USA. [3]Department of Pharmacology and Toxicology, Virginia Commonwealth University, Richmond, VA, USA. ✉e-mail: constantinople@nyu.edu

known to influence response vigor[20–23]. Therefore, RPEs in the NAcc, in addition to training animals to take reward-maximizing actions, likely influence whether actions in a state will be more or less vigorous[7]. The fact that exogenous 17β-estradiol influences dopamine release in the NAcc[4,24] suggests that endogenous fluctuations could influence RPEs for reinforcement learning. However, how behaviors unrelated to reproduction, such as reward-driven learning, respond to endogenous fluctuations in 17β-estradiol is unresolved.

17β-estradiol is expressed at low levels in brain tissue[25], and spatiotemporal dynamics of endogenous release are unknown. This makes it difficult to interpret exogenous application of synthetic estrogens as reflecting endogenous dynamics. To relate physiological fluctuations of 17β-estradiol to behavior and neural dynamics, we measured behavior and dopamine release in the NAcc over the reproductive cycle and suppressed 17β-estradiol's function in the midbrain by knocking down estrogen receptors. Challenges of this approach include the fact that certain reproductive stages are only expressed briefly (for example, proestrus lasts for 12 h once every 4–5 days), rats enter reproductive senescence unpredictably, and the likelihood of senescence increases with age, which becomes limiting for complex tasks with long training times. We overcame these challenges using a brute-force high-throughput behavioral training approach to train hundreds of rats on a rich decision-making task.

We find that rats update their response vigor on a trial-by-trial basis based on dopamine RPEs in the NAcc. Rats exhibit enhanced learning when endogenous 17β-estradiol is high, and estrogen receptors in the midbrain causally promote reinforcement learning. We show that following endogenous increases in 17β-estradiol, dopamine transporters are expressed at lower levels in the NAcc, and dopamine RPEs are enhanced. This reveals new roles for both 17β-estradiol and dopamine reuptake mechanisms in shaping RPEs for reinforcement learning.

Results

Estrogen levels predict enhanced learning

We trained rats on a self-paced temporal wagering task that manipulated reward expectations by varying reward magnitudes over blocks of trials[23] (Fig. 1a,b). Rats initiated trials by poking their nose in a center port. This triggered an auditory cue, the frequency of which indicated the volume of water reward offered on that trial (4, 8, 16, 32 or 64 μl). The reward was assigned to one of two side ports, indicated by a light-emitting diode (LED). The rat could wait for an uncued and unpredictable delay to obtain the reward, or at any time could opt out by poking in the other side port to start a new trial. On 15–25% of trials, rewards were withheld to assess how long rats were willing to wait for them (catch trials).

To manipulate reward expectations, rats experienced uncued blocks of trials in which they were presented with low (4, 8 or 16 μl) or high (16, 32 or 64 μl) reward volumes (16 μl was in both). These were interleaved with mixed blocks that offered all rewards (Fig. 1b). To control for potential trends due to satiety across the session, we regressed out initiation times against trial number. Detrended initiation times were faster in high blocks compared to low blocks (block sensitivity; Fig. 1c). Previous work has suggested that this pattern optimally balances the costs of vigor against the benefits of harvesting reward[26]. The sensitivity of initiation times to the reward blocks was robust across 303 male and female rats (Fig. 1d).

We previously found that the decision of when to initiate a trial was governed by a reinforcement learning algorithm that updated the value of the environment or state using RPEs[23,27]. In actor–critic models, the critic learns the value of the state (V), or the discounted sum of future expected rewards (R) in that state, and updates state values by the following equation: $V_{t+1} = V_t + \alpha(R_t − V_t)$, where $R_t − V_t$ is the RPE (δ), and α is the learning rate (Fig. 1e). This algorithm, also called a delta rule, computes the average of previous rewards, with most recent rewards having the strongest impact. Indeed, regressing trial initiation

times against previous reward offers yielded negative coefficients that declined exponentially[11], with more recent rewards having a stronger influence on behavior (Fig. 1f). When a rat received a large reward, it initiated the next trial more quickly (that is, rewards and initiation times were inversely proportional), consistent with positive RPEs increasing state values. The exponentially decaying influence of previous rewards on value estimates is a hallmark of reinforcement learning algorithms, including the delta rule. Collectively, this suggests that rats estimate their future expected rewards by integrating over past rewards via a delta rule, and initiate trials more quickly when they anticipate higher future expected rewards.

To determine whether hormonal fluctuations relate to reinforcement learning, we tracked female rats' reproductive or estrous cycles with vaginal cytology (Extended Data Fig. 1a). There are four stages of estrous, each lasting ~12–48 h. 17β-Estradiol is low during metestrus and diestrus and peaks during proestrus, which is followed by estrus, when ovulation occurs (Extended Data Fig. 1b). Serum 17β-estradiol expression patterns measured with an enzyme-linked immunosorbent assay (ELISA) validated stages identified by vaginal cytology (Extended Data Fig. 1c,d).

When female rats were in proestrus, the stage where 17β-estradiol and progesterone peak, their initiation times were more sensitive to the reward blocks compared to when those same rats were in diestrus, when 17β-estradiol and progesterone are low (Fig. 1g,h and Extended Data Fig. 1e). Detrended initiation times showed stronger (weaker) modulation of initiation times following block transitions in proestrus (diestrus; Fig. 1i). The difference between trial initiation times in low and high blocks was enhanced in proestrus and estrus, when ovulation occurs (Extended Data Fig. 1c,e). In diestrus, females showed block sensitivity that was closer to males (Fig. 1g,h).

Rats also adjusted how long they were willing to wait for rewards in each block[23]. There was no effect of the estrous cycle on how long they were willing to wait for rewards (Extended Data Fig. 2). Moreover, we previously found that the wait times are governed by a different behavioral strategy that uses state inference, and that is distinct from the trial-by-trial reinforcement learning strategy that governs initiation times[23]. Because the initiation times reflect a reinforcement learning process that we hypothesize relies on dopaminergic RPEs, and they are hormonally modulated, we focused on that aspect of behavior in this study.

To assess whether hormonal fluctuations modulate trial-by-trial learning at a finer timescale than reward blocks, we regressed detrended trial initiation times in mixed blocks in proestrus and diestrus against previous rewards. In proestrus, detrended initiation times were more strongly influenced by the reward on the most recent trial compared to diestrus and to males (Fig. 1j,k). The increased weight of the previous reward correlated with increased behavioral sensitivity to blocks, suggesting that these metrics of learning, although at different timescales, are related (Extended Data Fig. 1f). We next regressed detrended initiation times against the previous reward, but this time included stage as a categorical variable. Across rats, we found that there was a significant interaction between stage and the previous reward that increased the influence of the previous reward on behavior in proestrus ($P = 8.43 × 10^{−5}$).

We measured serum levels of 17β-estradiol with an ELISA in 17 rats performing the task across the estrous cycle. We first validated that this subset exhibited the behavioral effects we identified in the larger population (Extended Data Fig. 3a–c). We observed a significant correlation between serum 17β-estradiol and session-by-session learning across all stages of the cycle (Fig. 1l and Methods).

We next tested whether variability in circulating 17β-estradiol could explain variability in the degree to which the estrous cycle modulated behavior across rats (change in block sensitivity between proestrus and diestrus; Fig. 1h). Using 11 rats, we first validated that 17β-estradiol was indeed higher in proestrus during these training
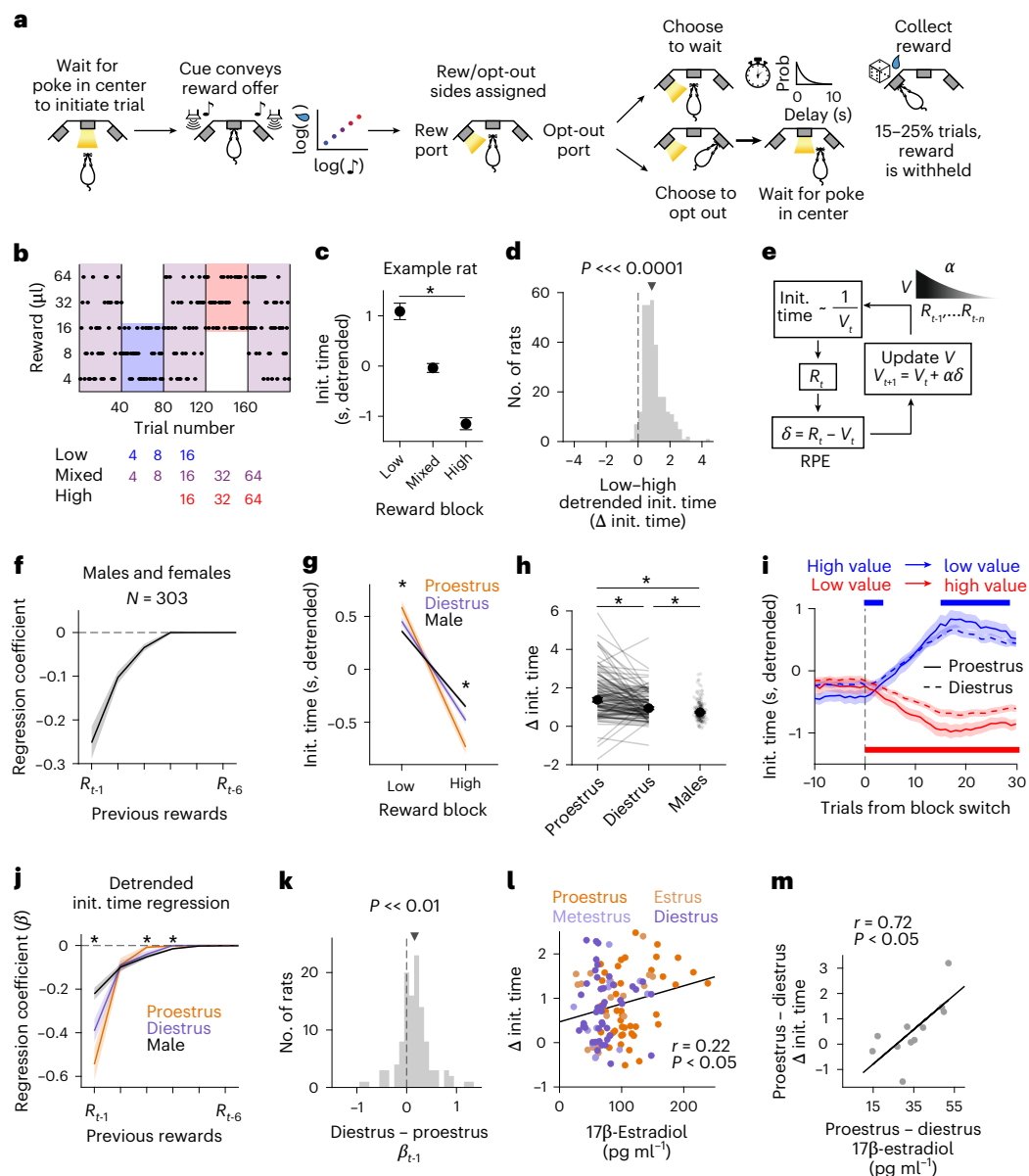
Fig. 1 | Rats' trial initiation times are modulated by the value of the state and estrous stage. a, Schematic of behavioral paradigm. Rew, reward; Prob, probability. **b**, Block structure of task with an example session (top) and reward distributions in each block (bottom). **c**, Mean detrended trial initiation times across blocks for one example rat (106 sessions). Initiation (Init.) times in low and high blocks are significantly different; $P \lll 0.0001$, two-sided Wilcoxon rank-sum test; error bars are confidence intervals. **d**, Sensitivity of detrended initiation times to blocks (low–high block) across the population is significantly different from zero; one-sided Wilcoxon signed-rank test, $N = 303$. **e**, Schematic depicting reinforcement learning model. Initiation times are inversely proportional to state value (V_t) on trial (t), which gets updated by an RPE (δ) when a reward is offered on that trial (R_t). A learning rate (α) determines how much to weigh the previous rewards (R_{t-n}) in the estimation of state values. **f**, Median regression coefficients of detrended trial initiation times as a function of rewards on previous trials during mixed blocks for all rats. **g**, Median detrended initiation times for low and high blocks, $N = 118$ female rats in proestrus and diestrus and $N = 185$ male rats; Kruskal–Wallis test $P = 1.04 \times 10^{-6}$ for low blocks and $P = 3.96 \times 10^{-14}$ for high blocks, two-sided Wilcoxon signed-rank test $P = 1.38 \times 10^{-5}$ and $d = 0.38$ for low blocks proestrus versus diestrus, $P = 5.94 \times 10^{-8}$ and $d = 0.48$ for high blocks, two-sided Wilcoxon rank-sum test $P = 2.91 \times 10^{-7}$ and $P = 1.57 \times 10^{-14}$ for low and high blocks proestrus versus males, and $P = 0.02$ and $P = 1.54 \times 10^{-4}$ for low and high blocks diestrus versus males. **h**, Median block sensitivity (low–high) of detrended trial initiation times. Two-sided Wilcoxon signed-rank test for

proestrus versus diestrus $P = 3.50 \times 10^{-8}$ and $d = 0.51$, two-sided Wilcoxon rank-sum test for proestrus versus males $P = 1.99 \times 10^{-12}$ and for diestrus versus males $P = 1.27 \times 10^{-3}$. **i**, Median detrended initiation times at transitions from relatively high-to-low value (blue) or low-to-high value (red) blocks have greater sensitivity to the change in value in proestrus than diestrus, $N = 118$. Causal filter spanning 15 trials; bars represent $P < 0.05$ using a two-sided Wilcoxon signed-rank test (actual P values listed in Supplementary Table 1). **j**, Mean regression coefficients are stronger in proestrus, one-way ANOVA $P = 2.42 \times 10^{-18}$ for one trial back ($d = 0.47$ for proestrus versus diestrus), $P = 0.85$ for two trials back, $P = 1.07 \times 10^{-7}$ for three trials back, $P = 9.23 \times 10^{-9}$ for four trials back, $P = 0.02$ for five trials back and $P = 0.04$ for six trials back). **k**, The distribution of the differences between regression coefficients of one trial back for proestrus and diestrus is significantly different from zero; one-sided Wilcoxon signed-rank test $P = 3.91 \times 10^{-9}$. Coefficients are more (less) negative in proestrus (diestrus). **l**, Block sensitivity is significantly correlated with serum 17β-estradiol levels across the cycle with $N = 17$ rats and 116 sessions; two-sided Pearson correlation $P = 0.018$. **m**, The differences in block sensitivity for proestrus and diestrus are significantly correlated with the change in 17β-estradiol levels on sampled sessions; two-sided Pearson correlation $P = 0.013$. $*P = 0.01$, $P < 0.01$, arrows are medians and error bars are all ±s.e.m. unless otherwise noted. Data in **j** and **k** only include rats with greater initiation times in low compared to high blocks ($N = 114$ female rats in proestrus and diestrus and $N = 179$ male rats).

sessions (Extended Data Fig. 1d). We uncovered a strong and significant correlation between changes in circulating levels of 17β-estradiol and hormonal modulation of block sensitivity, where rats with greater changes in 17β-estradiol had greater hormonal modulation of behavior (Fig. 1m). This suggests that the degree to which 17β-estradiol fluctuates over the cycle dictates whether the cycle influences learning and can explain why some rats show stronger or weaker modulation.

Collectively, these data show significant correlations between reinforcement learning and 17β-estradiol as it fluctuates over the reproductive cycle. While there is a drop in 17β-estradiol in estrus (Extended Data Fig. 1c), but still enhanced session-by-session learning in that stage (Extended Data Fig. 1e), this could reflect longer-lasting genomic effects following the transient 17β-estradiol surge in proestrus. We refrain from making claims about whether hormonal modulation of behavior, including enhanced sensitivity to previous rewards, is better or worse for performance, because the exact benefits (costs) that are maximized (minimized) are unobservable. Instead, we used a reinforcement learning framework to make precise descriptive statements about changes in learning over hormonal states.

Estrogen levels predict enhanced RPEs

We next sought to characterize the neural correlates of enhanced reinforcement learning in proestrus. To test whether physiological patterns of dopamine signaling encode an RPE (Fig. 2a) that is modulated by the estrous cycle, we virally expressed the optical dopamine sensor $GRAB_{DA}$ in the NAcc, and implanted a fiber optic over the injection site for fiber photometry (Fig. 2b). We used a static fluorophore (mCherry) to correct for brain motion (Methods and Extended Data Fig. 5e,f). We observed a strong phasic dopamine response when the rat heard the reward offer cue. Dopamine fluorescence at the offer cue scaled with the offered reward volume: there were positive phasic responses to larger rewards and negative responses to smaller rewards, consistent with an RPE (Fig. 2c,d and Extended Data Fig. 4). We interpret this result as being consistent with a temporal-difference learning algorithm[8,9], in which the RPE would propagate back to the earliest predictor of the value of the environment on each trial (here, the cue).

We next examined the 16 μl reward offers, which appeared in all reward blocks, but with a different value relative to the average reward in each block. In reinforcement learning algorithms, the value of the state (against which rewards are compared) is the discounted sum of future expected rewards in that state. In our task, the expected rewards varied with the reward blocks. The response to the auditory cue predicting 16 μl was positive in low blocks, when average reward was low, and negative in high blocks, when average reward was high, consistent with an RPE (Fig. 2a,e,f). Therefore, we focused subsequent analyses on the phasic dopamine response at the time of the offer cue.

We found that phasic dopamine release was more strongly modulated by reward states in proestrus compared to diestrus (Fig. 2g,h). Similar to what was observed behaviorally, in proestrus, the phasic dopamine response to the offer cue was more sensitive to the reward blocks (Fig. 2g). To quantify this effect, for each rat, we calculated the area under the curve (AUC) of the dopamine response in the 500 ms following the offer cue, for low and high blocks. For all rats, there was a larger difference in these values in proestrus (Fig. 2h). Furthermore, serum 17β-estradiol predicted variability in dopamine block sensitivity across sessions (Fig. 2i), establishing a significant correlation between 17β-estradiol and dopamine signaling. Enhanced dopamine sensitivity to reward states was driven by enhanced dopamine transients for cues predicting the largest rewards (32 μl and 64 μl) in proestrus (Fig. 2j,k and Extended Data Fig. 5a,b).

Dopamine release at recording sites dorsal to the NAcc in the caudate putamen also exhibited hallmarks of RPE encoding during the offer cue (Extended Data Fig. 6). However, dopamine at these sites was unaffected by the estrous cycle, suggesting that hormonal modulation of dopamine release may be specific to the NAcc.

Consistent with previous studies[28], we recently found that dopamine after the delay period exhibited dynamics that may reflect beliefs about hidden task states (whether the state was a rewarded or unrewarded trial)[27]. Dopamine aligned to the delay start exhibited negative ramps that have been interpreted as moment-by-moment negative RPEs as the animals waited without receiving a reward[14] (Extended Data Fig. 7a). The magnitude of phasic dopamine aligned to the reward cue, signaling the end of the delay, scaled with delay length (Extended Data Fig. 7b). This pattern requires additional knowledge about the task structure as the probability of being in an unrewarded trial increases over time, so the reward cue is more unexpected given beliefs about the trial[28]. This RPE was also larger in proestrus compared to diestrus (Extended Data Fig. 7c,d), especially at longer delays, suggesting that RPEs reflecting state inference were also enhanced during proestrus. Notably, these RPEs cannot account for the modulation of initiation times by block over estrous as the delays did not change over blocks.

To test whether enhanced RPEs at the offer cue could account for enhanced reinforcement learning in proestrus, we used a reinforcement learning model that predicts rats' trial initiation times[23,27]. In the model, initiation times were inversely related to state value, which was estimated with a delta rule (Fig. 3a). The model captured qualitative aspects of rats' behavior, including sensitivity of initiation times to the reward blocks and previous reward offers (Fig. 3b,c). To determine when RPEs were most strongly encoded, we fit the model to each rat's initiation times. Then we used the best-fitting parameters for each rat to estimate their trial-by-trial RPEs (Methods). We regressed RPEs from the behavioral model against the dopamine AUC in sliding windows around each event (Fig. 3d). We found the strongest encoding of RPEs (that is, largest regression coefficients) at the time of the offer cue, when rats hear the tone indicating the reward offer, consistent with a temporal-difference learning algorithm.

To quantify encoding of RPEs in proestrus and diestrus, we used the behavioral model—fit to all stages of the estrous cycle—to generate trial-by-trial RPE estimates for each rat. We then separately computed the dopamine AUC for different RPE values in proestrus versus diestrus. We found that in proestrus, dopamine encodes RPEs with a greater dynamic range. Moreover, there is a significant increase in the amplitude of the phasic dopamine response for large, positive RPEs (Fig. 3e; see also Fig. 2g,h,j,k and Extended Data Fig. 5a–c).

To account for these apparent effects on large positive RPEs, we modified the behavioral model to include a multiplicative gain (ϕ) on positive RPEs above the 90th percentile (Fig. 3f). This explicitly enhanced large positive RPEs; however, because these larger RPEs increased state values, this manipulation also produced a modest increase in the dynamic range of negative RPEs (Fig. 3f,g), consistent with dopamine measurements (Fig. 3e). Adding this ϕ parameter recapitulated the enhanced behavioral sensitivity to reward blocks in proestrus (Fig. 3h; see also Fig. 1h). Moreover, it changed the impact and time constant with which previous rewards influenced the model's behavior, and increased behavioral modulation at block transitions, consistent with behavior in proestrus compared to diestrus (Fig. 3i,j; see also Fig. 1i,j). Alternatively, an enhanced learning rate parameter could in principle increase the impact of previous rewards on behavior. However, an increased learning rate does not predict any change in the relationship between the current reward offer and the RPE. In proestrus, dopamine showed enhanced responses to the current reward offer, which was captured by changing the ϕ parameter in the model, but not the learning rate, suggesting that rats in proestrus specifically exhibit larger positive RPEs (Fig. 3k–m).

Collectively, these data show that dopamine dynamics at the time of the offer cue encoded an RPE. Dopamine responses in the NAcc at this time point scaled with the offered reward (Fig. 2c,d), were inversely proportional to reward expectations (Fig. 2e,f), correlated with RPEs estimated from the behavioral model (Fig. 3d) and exhibited a greater
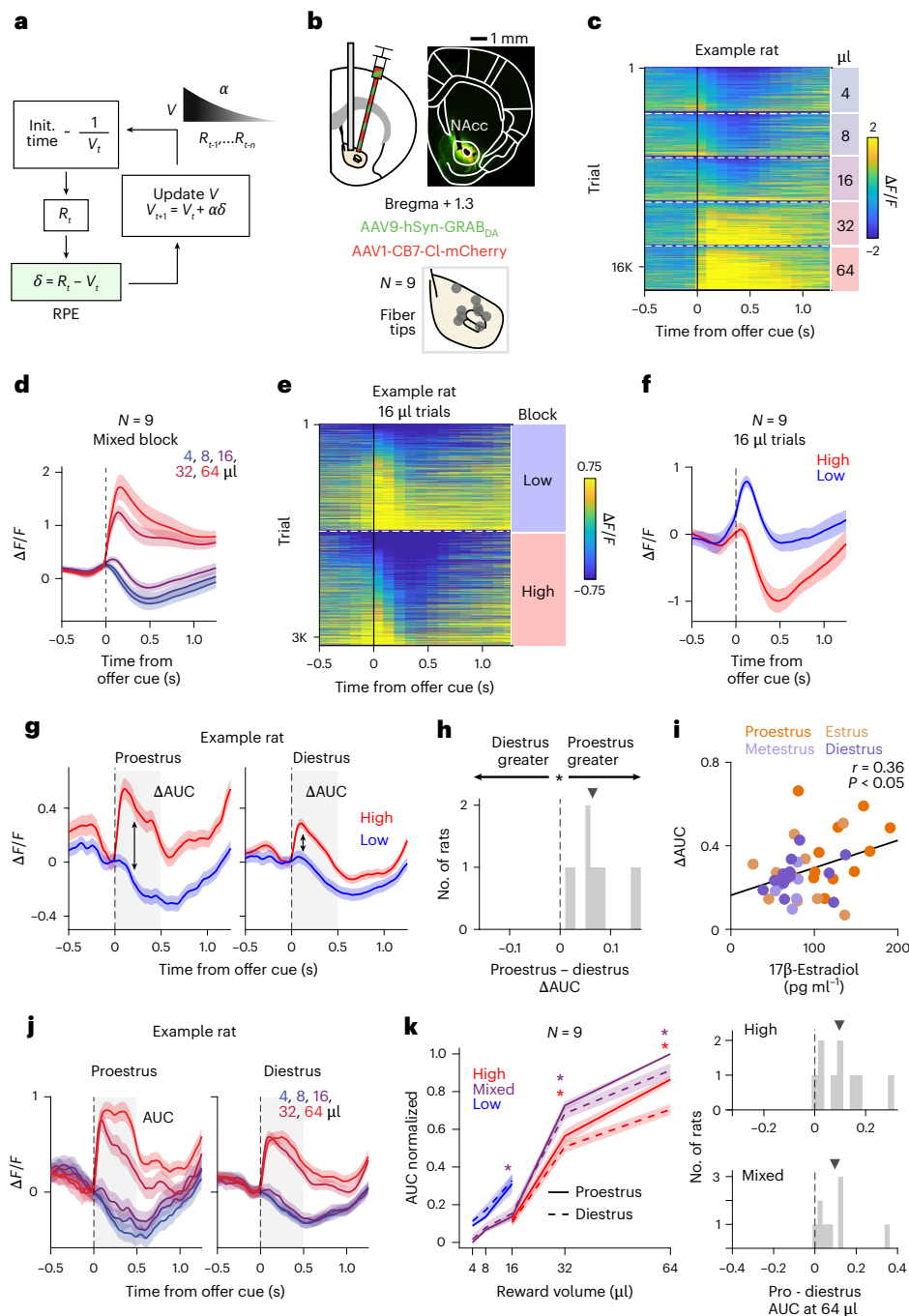
Fig. 2 | NAcc dopamine at reward offer cue encodes RPEs with greater dynamic range in proestrus. a, Schematic depicting delta rule with RPE highlighted. **b**, NAcc GRAB$_{DA}$ and mCherry virus injections in female rats, with fibers. Top left: schematic and representative example (top right) of GRAB$_{DA}$ expression in green and mCherry in red. Bottom: fiber locations in the NAcc for nine female rats. **c**, Heat map of GRAB$_{DA}$ signal in response to reward offer cue across >16,000 trials recorded from an example rat during mixed blocks and sorted by reward volume and peak response in the first 500 ms following offer cue. Time bins are 100 ms. **d**, Average response to offer cue during mixed blocks, separated by reward volume. **e**, Heat map of GRAB$_{DA}$ signal in response to reward offer cue for 16 μl across >3,000 trials recorded from an example rat sorted by reward block and peak response in the first 500 ms following offer cue. **f**, Average response to 16 μl offer cue across the population, separated by reward block. **g**, Response to offer cue, separated by reward block and stage group for an example rat. Gray box represents the window used to calculate the change in area under the

curve (ΔAUC) 500 ms after the offer cue in high–low blocks in **h**. **h**, Histogram of stage effect (proestrus–diestrus) on ΔAUC for offer cue response over rats shows stronger block encoding in proestrus; two-sided Wilcoxon signed-rank test $P = 3.91 \times 10^{-3}$. Triangle is the median. **i**, Dopamine ΔAUC in high–low blocks, as a function of circulating levels of 17β-estradiol across 41 sessions; Pearson correlation coefficient $P = 0.019$, $N = 5$. **j**, Response to offer cue, separated by reward volume and stage group for an example rat. **k**, Left: median population response to each reward volume, separated by reward block and min–max normalized. Right: median AUC of individual rat responses for the largest reward volume for high and mixed blocks. Triangles are the medians. AUC is larger in proestrus than diestrus; two-sided Wilcoxon signed-rank test during mixed blocks at 16 μl ($P = 0.039$, $d = 0.19$), 32 μl ($P = 0.02$, $d = 0.49$) and 64 μl ($P = 0.0039$, $d = 0.89$) and during high blocks at 32 μl ($P = 0.02$, $d = 0.53$) and 64 μl ($P = 0.0039$, $d = 1.64$). *$P < 0.05$, all error bars are ± s.e.m. In **g**–**k**, data are baseline-corrected using 50 ms before the offer cue.
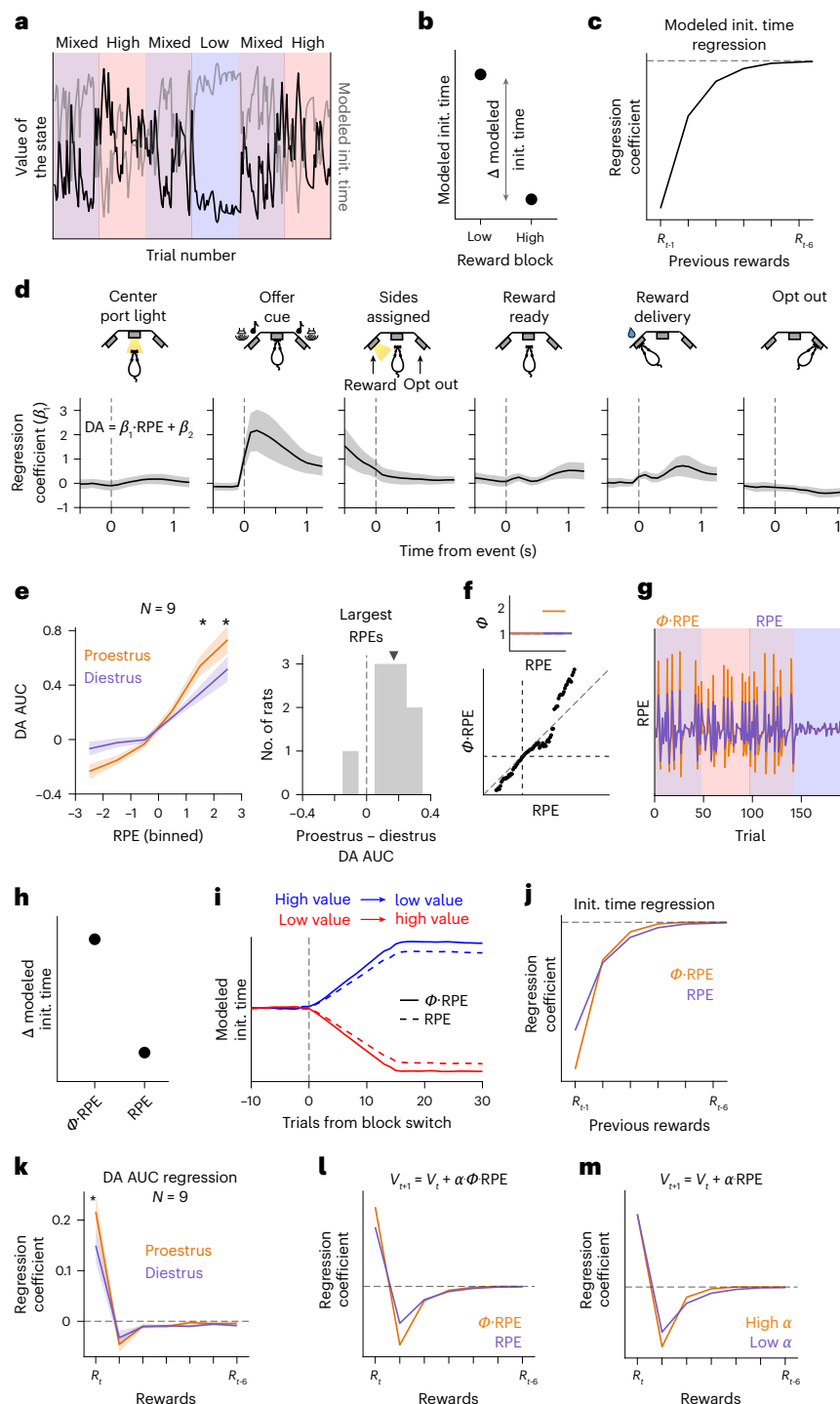
Fig. 3 | Multiplicative gain on large RPEs in a reinforcement learning model captures enhanced learning from rewards in proestrus. a, Simulation of the value of the state across reward blocks (blue, low; purple, mixed; red, high) using a reinforcement learning model and predicted initiation times for one session. **b**, Model-predicted initiation times by block. **c**, Regression of model-predicted initiation times against rewards on previous trials during mixed blocks. **d**, Regression coefficients of model-predicted RPEs (estimated from each rat's best-fitting model parameters fit to the last ten trials of the block) as a function of dopamine in each time bin (100 ms) aligned to task events, with the strongest encoding at the offer cue. Error bars are the mean ± s.e.m. **e**, Left: median dopamine AUC at the offer cue as a function of model-predicted RPEs is greater for proestrus than diestrus for the two largest positive RPE bins; Wilcoxon

signed-rank test $P = 0.012$, $d = 0.57$ and $P = 0.012$, $d = 0.55$. Right: difference in AUC for the largest RPE bin by rat. Triangle is the median; error bars are the median ± s.e.m. **f**, Multiplicative gain (ϕ, 1.75) was applied to large RPEs (RPEs > 90th percentile). **g**, Increasing large RPEs enhanced trial-by-trial RPEs, mimicking proestrus. **h–j**, Increasing ϕ on large RPEs captured enhanced learning seen in proestrus, including greater block sensitivity (**h**), more pronounced behavioral changes at block transitions (**i**) and greater impact of previous rewards on initiation time (**j**). Error bars in **i** are the median ± s.e.m. **k**, Rewards on the current trial have a significantly larger impact on dopamine in proestrus compared to diestrus; Wilcoxon signed-rank test $P = 0.02$, $d = 0.21$. Error bars are the median ± s.e.m. **l,m**, Enhancing ϕ captured the increased regression coefficient on the current trial (**l**), while enhancing the learning rate, α, did not (**m**). DA, dopamine. *$P < 0.05$.
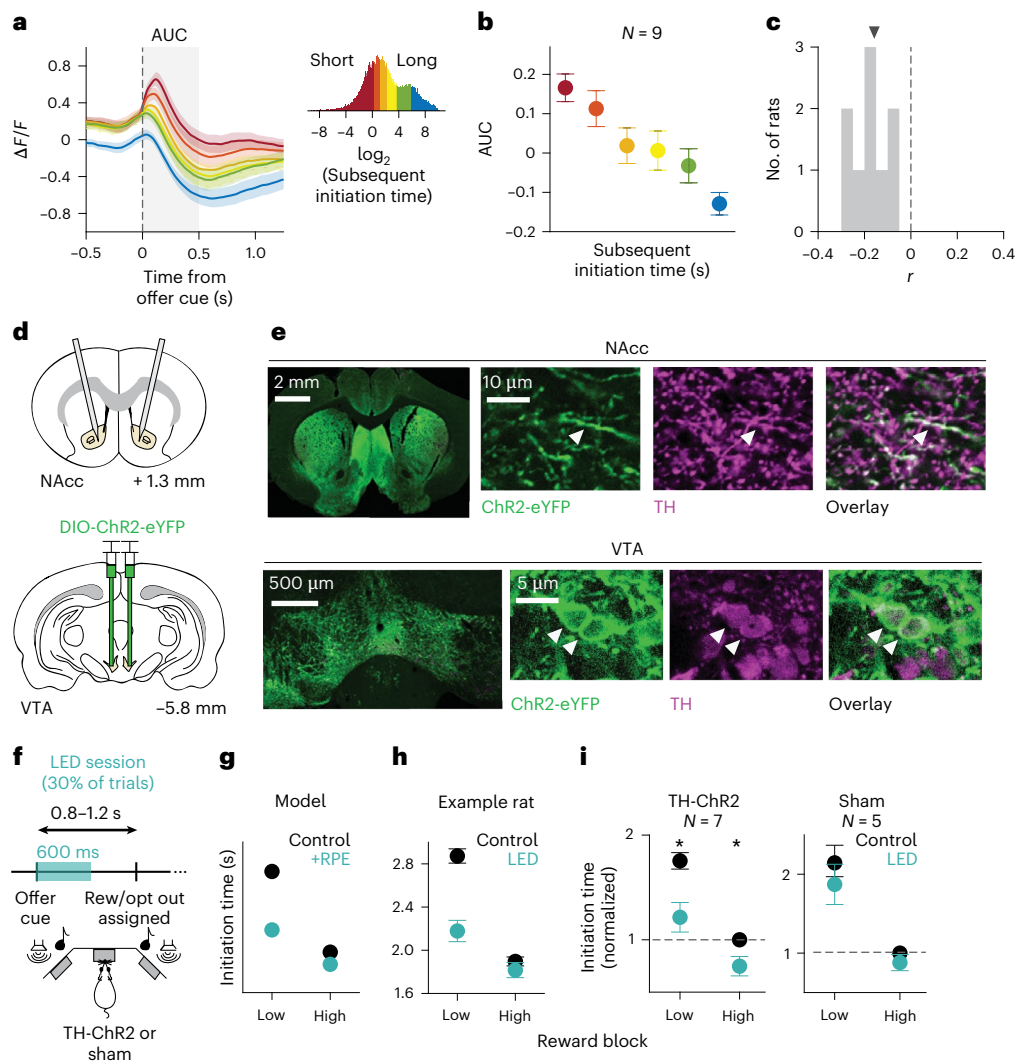
Fig. 4 | Phasic dopamine release in the NAcc controls trial initiation time, consistent with an RPE that updates state values. a, Dopamine signal aligned to the offer cue split by the subsequent trial initiation time (<0.2, 0.4, 0.9, 2.5, 12 and >12 s), N = 9 rats. Inset histogram shows how $\log_2$(initiation times) were binned. **b**, Mean AUC of curves in a 500-ms window across nine rats shown in **a**. **c**, Two-sided Pearson correlation of the AUC and subsequent initiation time over rats; all $P < 0.01$. Actual P values are available in Supplementary Table 1. Triangle is the median. **d**, Schematic of the strategy used to specifically stimulate TH⁺ VTA terminals expressing ChR2 in the NAcc. **e**, ChR2-YFP colocalization with TH in VTA cell bodies and NAcc axons (×20 and ×63). Arrowheads point to labeled

NAcc axons or VTA cell bodies. Example image is one of 90. **f**, TH⁺ axon terminals in the NAcc were stimulated with 465-nm light during the offer cue on 30% of trials. **g**, Simulated initiation times with and without an additive RPE gain on 30% of trials, using the behavioral model. **h**, Mean trial initiation times for an example TH-Cre rat for 10,892 trials with optogenetic stimulation and 17,206 trials without (control). **i**, Mean trial initiation times across the population normalized to high blocks in control sessions for rats expressing ChR2 in TH⁺ axons (left) and sham controls (right); two-sided Wilcoxon signed-rank test $P = 0.02$, $d = 1.47$ for low blocks and $P = 0.03$, $d = 1.35$ for high blocks. All error bars are the mean ± s.e.m., *$P < 0.05$.

dynamic range in proestrus compared to diestrus (Figs. 2h,k and 3e). Hormonal modulation of female rats' reinforcement learning could be accounted for by a multiplicative gain on large positive RPEs, which was consistent with dopamine measurements in the NAcc.

NAcc dopamine controls initiation times

According to the behavioral model, dopamine RPEs update the value of the state, which determines the vigor with which animals initiate subsequent trials. To test this hypothesis, we evaluated dopamine on trials with different subsequent initiation times. In previous work, we found that initiation times were most strongly governed by RPEs following unrewarded trials[23], so we restricted this analysis to those trials. Indeed, trials in which subsequent initiation times were fast showed larger phasic dopamine responses, and trials with long subsequent initiation times showed apparent dips in dopamine (Fig. 4a). There was a significant negative correlation between the subsequent initiation

time on trial $t + 1$, and the amplitude of the dopamine on trial t, in all rats (Fig. 4b,c). In other words, the phasic dopamine response at the offer cue predicted the rats' initiation time on the next trial, consistent with dopamine at this event acting as an RPE to update the value of the state for response vigor.

To test whether dopamine at this time point causally influenced initiation times, we used optogenetics to activate dopamine terminals in the NAcc. We expressed Cre-dependent excitatory opsins (adeno-associated virus 9 (AAV9)-DIO-ChR2.0) bilaterally in the VTA and substantia nigra of tyrosine hydroxylase (TH)-Cre rats and rats injected with a TH-Cre virus, and bilaterally implanted fiber optics in the NAcc (Fig. 4d,e). We histologically verified that ChR2 expression was largely restricted to neurons that were TH immunoreactive (Fig. 4e and Extended Data Fig. 8). On 30% of trials, we optogenetically stimulated midbrain dopamine terminals in the NAcc triggered when rats heard the tone indicating the reward offer.

We reasoned that this manipulation was similar to providing an additive gain to RPEs because it would add an (approximately) constant amount of dopamine on stimulation trials. Simulating an additive RPE gain in the behavioral model predicted that rats should exhibit faster trial initiation times because positive prediction errors would increase state values (Fig. 4g). The model predicted a larger effect in low blocks, because it implemented a floor effect to capture the fact that rats can only initiate trials so quickly in high blocks. Indeed, we found that dopamine terminal stimulation decreased initiation times (Fig. 4h,i). There was no effect on trial initiation times in sham animals that did not express light-sensitive opsins (Fig. 4i). There was also no effect on wait times (two-way ANOVA, $F = 0.24$ and $P = 0.91$ for an interaction between reward and session type for TH-Cre rats and $F = 0.207$ and $P = 0.103$ for sham rats), further showing that dopamine at reward offer cue encodes an RPE that is selectively used to guide initiation times. Collectively, these results support the hypothesis that dopamine in the NAcc at the time of the offer cue acts as an RPE to update state values for subsequent initiation times. We note that while we used continuous light stimulation, our results are consistent with optogenetic studies that used pulse-based stimulation[17,29,30].

Dopamine transporters vary with estrogen levels

Estrogen receptors can produce longer-term genomic changes (that is, synthesis of new proteins that can last for hours to days), or activate rapid signaling cascades whose effects persist on shorter timescales (that is, for minutes or hours)[3]. To test for changes in protein expression, we performed quantitative mass spectrometry of the NAcc in diestrus, proestrus and estrus ($N = 5$–6 rats in each group). A total of 238 proteins were significantly decreased in expression in proestrus or estrus compared to diestrus, and 258 were increased. Gene Ontology analysis identified several proteins involved in dopamine signaling that were differentially expressed in these stages (Methods and Extended Data Fig. 9). However, the only differentially expressed proteins that could account for the observed increase in phasic dopamine in proestrus and estrus were proteins involved in dopamine reuptake, which were expressed at lower levels in these stages (Fig. 5a and Extended Data Fig. 9). The dopamine transporter (DAT or Slc6a3) controls dopamine reuptake in the NAcc, and was expressed at significantly lower levels in proestrus compared to diestrus. The serotonin transporter (SERT or Slc6a4) is also capable of robust dopamine reuptake[31], and was expressed at lower levels in proestrus and estrus compared to diestrus. Smpd3 (sphingomyelin phosphodiesterase 3) triggers fusion of DAT to the plasma membrane, and has been shown to make dopamine reuptake more efficacious[32,33]. Smpd3 was decreased in expression in estrus, suggesting a lower density of DAT in the plasma membrane during estrus, consistent with effects of 17β-estradiol applications on DAT localization in vitro[33]. Reduced plasmalemmal DAT and SERT should, in principle, lead to higher levels of extracellular dopamine.

We verified that rats exhibited reduced levels of DAT and SERT in proestrus and estrus, using a different assay, immunohistochemistry, and a separate cohort of rats in diestrus, proestrus and estrus ($N = 3$ per group). All sections were co-incubated in the same solutions simultaneously, to ensure that differences in fluorescence did not reflect subtle differences in tissue processing. We quantified fluorescently labeled pixel area from confocal photomicrographs, where fluorescence indicated DAT or SERT immunoreactivity (Fig. 5b). Consistent with the proteomics results of reduced DAT and SERT expression in these stages, we found that significantly fewer pixels showed DAT and SERT immunoreactivity in proestrus and estrus compared to diestrus (Fig. 5c,d). These results, therefore, held across different assays (proteomics and immunohistochemistry) and cohorts of rats.

Smpd3 is implicated in triggering fusion of DAT to the plasma membrane, and was also decreased in expression in estrus. To test whether there was also a change in subcellular localization of DAT in estrus, we characterized DAT localization at the ultrastructural level.

Individual DAT particles were labeled with silver-intensified gold (SIG) and visualized with electron microscopy. We quantified the fraction of DAT particles that were continuous with the membrane (membranous) versus discontinuous (intracellular), and also measured the distance from the center of the intensified gold particle to the plasma membrane. There was no significant difference in the fraction of DAT particles that were membranous versus intracellular, nor was there a difference in DAT particles' proximity to the membrane between estrus and diestrus (Fig. 5e,f). Therefore, reduced levels of DAT and SERT expression, but not differences in subcellular localization of DAT, likely account for enhanced dopamine RPEs and learning in proestrus and estrus.

How could changes in reuptake, which ostensibly would result in elevated levels of extracellular dopamine, account for the enhanced phasic responses we observed? We created a computational model in which action potentials of dopamine neurons caused instantaneous increases in dopamine concentration, which decayed with a rate that reflected reuptake dynamics, unless another action potential caused another instantaneous increase. The fluorescence signal at each time step was treated as a noisy observation of the underlying dopamine concentration[34]. The model assumed that the Poisson spike rate of the dopamine neurons increased or decreased (to zero) for positive or negative RPEs, respectively. When the dopamine concentration decayed more slowly, to simulate reduced reuptake, this increased the likelihood that dopamine from consecutive action potentials would temporally summate before decaying back to baseline. Averaged over trials, this temporal summation mechanism produced larger positive phasic responses (Fig. 5g–i), consistent with what we observed by measuring dopamine in proestrus at the offer cue (Figs. 2j and 3e) and after the reward delay (Extended Data Fig. 7d). Moreover, increasing temporal summation modestly elevated baseline levels of dopamine, producing a greater dynamic range for negative RPEs, which is normally thought to be limited by the low background firing rates of dopamine neurons[11]. These results are consistent with previous studies showing that inhibition of DAT increases both tonic and phasic dopamine[35,36]. In our simulation, reduced reuptake produced the strongest enhancement of the largest phasic DA responses, similar to our empirical observation that large positive RPEs were most strongly enhanced in proestrus. Therefore, decreased expression of DAT and SERT could produce a greater dynamic range of phasic dopaminergic RPEs by promoting temporal summation of dopamine from consecutive action potentials.

Midbrain ERα knockdown impairs learning

Many hormones fluctuate over the estrous cycle. We next sought to test whether estrogenic signaling in the midbrain, specifically, influenced reinforcement learning. Expression of DAT in the NAcc is restricted to dopaminergic axons from the midbrain. Therefore, we hypothesized that estrogen receptors, which can regulate protein expression via transcription and other mechanisms when bound to 17β-estradiol[1–3], might modulate learning by influencing expression of DAT in dopaminergic neurons. We virally expressed short-hairpin RNA (shRNA), labeled with eYFP, to knockdown expression of the estrogen receptor ERα bilaterally in the VTA (sh*Esr1*-eYFP; Fig. 6a,b). We used either a virus that was inducible by doxycycline (dox) in food and water (U6/tetO-sh*Esr1*) or one that was constitutively active (lenti-sh*Esr1*; Methods). For each virus, we used fluorescence in situ hybridization (FISH) of fixed tissue sections to measure levels of labeled *Esr1* (ERα transcript). Quantification of labeled probes showed a significant reduction in *Esr1* in rats virally expressing shRNAs (Fig. 6c,d).

Genetic suppression of estrogen receptors in the midbrain significantly reduced rats' behavioral sensitivity to the reward blocks (Fig. 6e). This is consistent with the finding that block sensitivity was correlated with 17β-estradiol over the estrous cycle. In fact, the reduction in sensitivity is comparable to how sensitivity is reduced in males
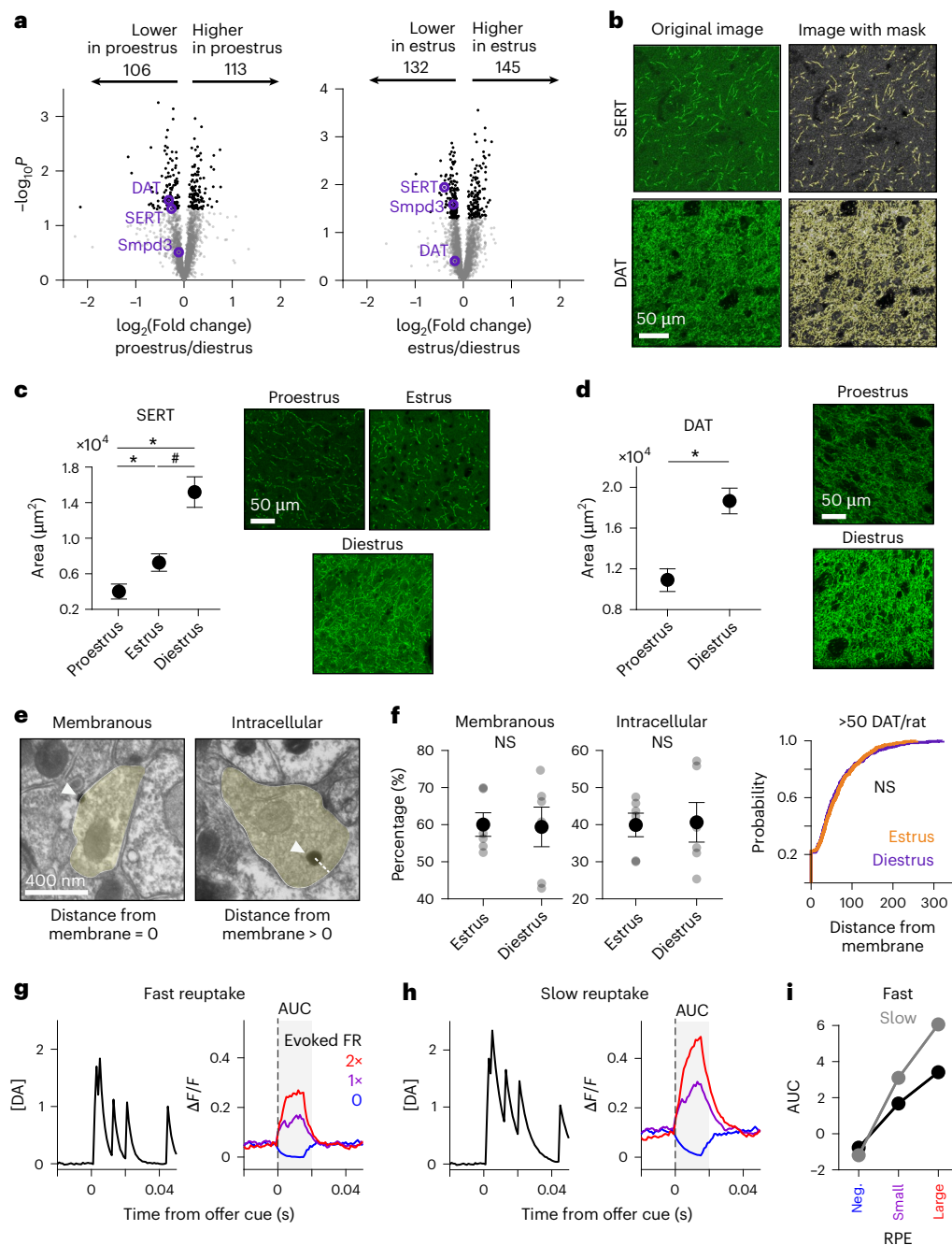
Fig. 5 | Dopamine reuptake proteins are decreased in expression in proestrus and estrus. a, Volcano plot depicting significantly differentially expressed proteins (black), *t*-test and a Benjamini–Hochberg false discovery rate of 5%. Proteins related to dopamine reuptake that were decreased in expression in proestrus and estrus are highlighted in purple. **b**, Example images from immunohistochemical quantification of SERT (36–54 micrographs per stage) and DAT expression (50–54 micrographs per stage) at ×63 with overlaid mask identifying fluorescence pixels above a threshold. **c**, Area of mask segmentation for SERT was highest in diestrus micrographs; Kruskal–Wallis test for group effect $P = 4.89 \times 10^{-3}$ and Wilcoxon rank-sum post hoc tests $P = 5.17 \times 10^{-3}$, $d = 2.2$ for proestrus versus diestrus, $P = 0.052$, $d = 1.13$ for estrus versus diestrus and $P = 0.024$, $d = 0.64$ for proestrus versus estrus. Example micrographs from each stage. **d**, DAT was also reduced in proestrus compared to diestrus micrographs;

two-sided Wilcoxon rank-sum test $P = 3.66 \times 10^{-5}$, $d = 0.92$. **e**, Example of DAT labeled with SIG and imaged with electron microscopy. Arrowheads point to putative DAT proteins. Minimum distances between plasma membrane and center of DAT particle are shown with dashed lines. Axons are highlighted in yellow. **f**, Proportion of 895 membranous and intracellular DAT particles (left) and their distance from the membrane, expressed as a cumulative probability (right), quantified from electron microscopy images does not change by stage; Wilcoxon rank-sum $P = 1$ for membranous versus intracellular and Kolmogorov–Smirnov test $P = 0.91$. **g,h**, Simulation of dopamine concentrations and fluorescence following different evoked firing rates (FRs), with fast (**g**) or slow (**h**) reuptake dynamics. **i**, Quantification of AUC in **g** and **h**. All error bars are the mean ± s.e.m., *$P < 0.05$, #$P < 0.1$. NS, not significant.

and hormonally modulated females in diestrus compared to proestrus (Fig. 6f). Reducing the RPE gain parameter in the delta model recapitulated the effect of sh*Esr1* (Fig. 6g). We conclude that midbrain estrogen receptors promote reinforcement learning by increasing the impact

of previous rewards on behavior. We note that the effects of sh*Esr1* on RPEs are correlatively drawn through these effects on behavior.

In proestrus, rats also appeared to be less thirsty: their initiation times were significantly slower, they performed fewer behavioral
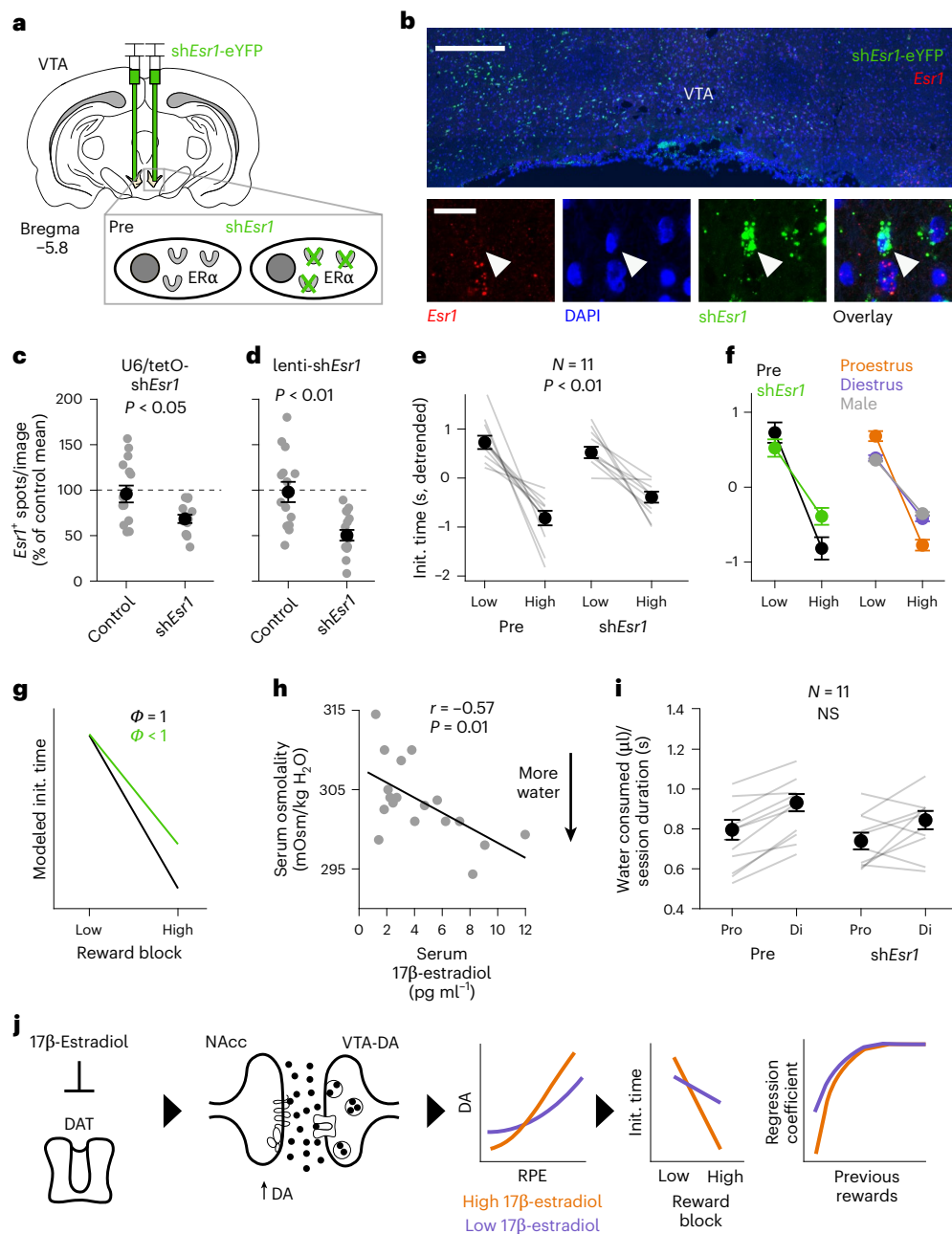
Fig. 6 | Midbrain ERα knockdown suppresses reinforcement learning.
a, sh*Esr1*-eYFP was injected bilaterally in the VTA. **b**, Expression of eYFP-labeled sh*Esr1* virus in the VTA (×20, top) after RNAscope. Scale bar, 500 μm. Expression of *Esr1* in eYFP-labeled cells compared to cells not expressing eYFP (×63, bottom). Scale bar, 20 μm. Arrowheads point to lack of *Esr1* expression in eYFP-labeled cells. **c,d**, Number of *Esr1*-labeled spots per photomicrograph of a *z*-stack taken from a rat injected with U6/tetO-sh*Esr1* (**c**) and lenti-sh*Esr1* (**d**) compared to not injected control tissue, expressed as a percentage of the control means, *N* = 14–15 micrographs. *Esr1* is suppressed with sh*Esr1* compared to controls; two-sided Wilcoxon rank-sum tests for U6/tetO-sh*Esr1* *P* = 0.014, *d* = 1.0 and lenti-sh*Esr1* *P* = 2.41 × 10⁻³, *d* = 1.15. **e**, Detrended initiation time across the population before and after sh*Esr1* injection for low and high blocks. Block sensitivity (low–high) was significantly reduced after sh*Esr1*; two-sided Wilcoxon signed-rank test *P* = 1.95 × 10⁻³, *d* = 0.75, *N* = 11 rats. **f**, Median change in block sensitivity is comparable to proestrus compared to diestrus and males in hormonally modulated rats; *N* = 120 rats. **g**, Reducing *ϕ* to below 1 replicates suppression of behavioral sensitivity with sh*Esr1*. **h**, Serum osmolality is significantly correlated with serum 17β-estradiol; two-sided Pearson correlation *P* = 0.013. Line of best fit overlaid (*N* = 18). **i**, Water consumed per session, controlling for the total amount of time the rat was allotted in the task, was still reduced in proestrus compared to diestrus, and this reduction was not affected by sh*Esr1*; *N* = 11 rats, two-sided Wilcoxon signed-rank test comparing difference in water consumption in proestrus and diestrus before injection to after sh*Esr1* *P* = 0.23 and *d* = −0.6. **j**, Schematic summarizing findings. 17β-Estradiol reduces expression of dopamine transporters. This increases the amount of dopamine in the extracellular space, allowing for a greater dynamic range of RPEs and enhanced learning from previous rewards in proestrus. All error bars are the median ± s.e.m., unless otherwise noted.

trials and drank less water, compared to diestrus (Extended Data Fig. 10a–c). Measurements of serum osmolality indicated reduced thirst in proestrus, and were strongly correlated with serum levels of 17β-estradiol (Extended Data Fig. 10d,e and Fig. 6h). Notably,

detrending the initiation times minimized satiety effects, and allowed us to resolve estrogenic influences on trial-by-trial reinforcement learning. However, because the ERα knockdown was specific to the VTA, this experiment provided an opportunity to directly test whether

estrogenic modulation of reinforcement learning occurs independent of systemic effects on fluid balance. We compared rats' water consumption before and after knockdown of ERα. Despite reducing behavioral sensitivity to the reward blocks, ERα knockdown had no effect on estrous-dependent water consumption (Fig. 6i). These results indicate that endogenous 17β-estradiol impacts reinforcement learning and fluid balance through dissociable mechanisms: the former relies on ERα in the VTA, but the latter does not.

Discussion

Using a task that provided a continuous behavioral measure of state values for reinforcement learning, we found that rats learn more from previous rewards following increases of endogenous 17β-estradiol, building on previous studies that found sex-dependent effects of reward on motivation[37]. We propose a computational and circuit-level mechanism that could result in midbrain 17β-estradiol-dependent enhanced learning: 17β-estradiol suppresses expression of DAT and SERT in the NAcc, leading to reduced reuptake of dopamine and a greater dynamic range of dopamine that enhances RPEs, which causally influence initiation times on subsequent trials (Fig. 6j). The enhancement is most pronounced for large positive RPEs, possibly due to increased occupancy of pools of reuptake proteins at high dopamine concentrations. Expression of DAT and SERT in the striatum is highly dynamic[38,39], with protein turnover occurring over hours to days[39–41], suggesting that plasmalemmal transporter expression could plausibly be reduced due to hormonal fluctuations in proestrus and remain reduced through estrus.

One key future question is whether 17β-estradiol modulates the firing rates of dopaminergic cells in the VTA, in addition to dopamine dynamics in the NAcc. While previous studies have examined estrous-dependent VTA dynamics in the context of drug exposure[42,43], it is unclear whether 17β-estradiol affects encoding of RPEs in the VTA during reinforcement learning. 17β-Estradiol could, in principle, act directly on VTA neurons[44,45], or have indirect effects by modulating afferent inputs[46]. Medium spiny neurons in the NAcc project back to the VTA[47], so it is possible that hormonal modulation of dopamine dynamics in the NAcc described here could indirectly influence VTA activity, via this projection.

Here, we find a new role for dopamine transporters in shaping RPEs for reinforcement learning. Drugs of abuse such as cocaine and methamphetamine act on DAT[48], and both DAT and SERT are the targets of pharmacological treatments for neuropsychiatric disorders including schizophrenia, depression and anxiety[49]. Computational psychiatry models have examined these disorders through the lens of reinforcement learning[50–52]. Our results provide a potential biological mechanism that bridges dopamine reuptake and reinforcement learning in health and disease, and might also underlie hormonal modulation of symptom severity in females[53–55]. The mechanisms by which ERα activation leads to reduced expression of dopamine reuptake proteins are important future directions. These could involve ERα-induced signal transduction pathways that control transcription, translation or protein degradation, or a mechanism involving other hormone receptors such as the progesterone receptor, a canonical target of ERα[56].

We used a reinforcement learning model to characterize dopamine RPEs over the reproductive cycle. However, recent findings that dopamine neurons exhibit heterogeneous responses[16,57,58] and respond to unexpected events independent of reward[59–61] have challenged the RPE hypothesis of dopamine. One proposal is that dopamine encodes not RPEs but generalized sensory prediction errors, and these prediction errors are vector-valued as opposed to scalar[59–61]. Sensory prediction errors would support learning about sensory associations (that is, for successor representation), with different prediction errors broadcast to distinct downstream circuits. In this scenario, the RPE would be one element of the vectorized prediction error[60,61]. Based on previous studies, it seems likely that the NAcc would receive the RPE from that

vectorized signal[9–17]. An alternative proposal has argued that dopamine neurons encode RPEs (as opposed to sensory prediction errors), but different neurons encode RPEs for distinct features of the state representation, producing heterogeneity across neurons that can include reward-independent responses[62]. In this framework, as in theories of distributional reinforcement learning[63], pooling the activity of many dopamine neurons by measuring release would approximate scalar RPE[62]. Therefore, our interpretation of dopamine release in the NAcc as encoding RPEs is consistent with any of these theories. While our data are also consistent with a proposal that dopamine signals prediction errors related to causal inference[64], this new theory is still being tested[65].

A deeper challenge to the RPE hypothesis has come from studies showing that optogenetically perturbing dopamine can influence learning of neutral sensory associations[66,67] without affecting conditioned responding[67], suggesting that dopamine might support valueless associative learning rather than value learning. However, we recently found that cholinergic dynamics determine whether dopamine supports striatal-dependent value learning[68]. Therefore, cholinergic dynamics in the striatum could have mitigated value learning in these studies, whereas sensory associations may reflect dopamine-dependent learning in other circuits. Other findings, including valence-independent dopamine responses[69], may be difficult to interpret without knowledge of concurrent striatal cholinergic dynamics, which may gate whether those signals act as RPEs[68].

There is a major gap in our understanding of how endogenous hormones interact with neuromodulatory systems during rich, quantifiable behaviors. We studied physiological hormonal fluctuations, as opposed to exogenous applications following ovariectomy, because the latter produces compensatory changes in the brain and body[70], and likely only crudely approximates endogenous dynamics. Our computational approach provided a principled framework for interpreting behavioral, physiological, optogenetic and molecular profiling data, and represents a new approach for studying endocrine systems in the brain. This approach allowed us to characterize hormonal modulation of dopamine at broad levels of description, and opens up new avenues of research of hormone–neuromodulatory interactions during complex cognitive behaviors.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-025-02104-z.

References

1. Marino, M., Galluzzo, P. & Ascenzi, P. Estrogen signaling multiple pathways to impact gene transcription. *Curr. Genom.* **7**, 497–508 (2006).
2. McEwen, B. S. & Milner, T. A. Understanding the broad influence of sex hormones and sex differences in the brain. *J. Neurosci. Res.* **95**, 24–39 (2017).
3. Vasudevan, N. & Pfaff, D. W. Non-genomic actions of estrogens and their interaction with genomic actions in the brain. *Front. Neuroendocrinol.* **29**, 238–257 (2008).
4. Di Paolo, T. Modulation of brain dopamine transmission by sex steroids. *Rev. Neurosci.* **5**, 27–42 (1994).
5. Yoest, K. E., Quigley, J. A. & Becker, J. B. Rapid effects of ovarian hormones in dorsal striatum and nucleus accumbens. *Horm. Behav.* **104**, 119–129 (2018).
6. Cox, J. & Witten, I. B. Striatal circuits for reward learning and decision-making. *Nat. Rev. Neurosci.* **20**, 482–494 (2019).
7. Floresco, S. B. The nucleus accumbens: an interface between cognition, emotion, and action. *Annu. Rev. Psychol.* **66**, 25–52 (2015).

8. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 2018).
9. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
10. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
11. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
12. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).
13. Day, J. J., Roitman, M. F., Wightman, R. M. & Carelli, R. M. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* **10**, 1020–1028 (2007).
14. Kim, H. R. et al. A unified framework for dopamine signals across timescales. *Cell* **183**, 1600–1616 (2020).
15. Steinberg, E. E. et al. A causal link between prediction errors, dopamine neurons and learning. *Nature Neurosci.* **16**, 966–973 (2013).
16. Parker, N. F. et al. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neurosci.* **19**, 845–854 (2016).
17. Hamid, A. A. et al. Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**, 117–126 (2016).
18. Joel, D., Niv, Y. & Ruppin, E. Actor–critic models of the basal ganglia: New anatomical and computational perspectives. *Neural Netw.* **15**, 535–547 (2002).
19. Chen, R. & Goldberg, J. H. Actor–critic reinforcement learning in the songbird. *Curr. Opin. Neurobiol.* **65**, 1–9 (2020).
20. Niv, Y., Daw, N. D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* **191**, 507–520 (2007).
21. Xu-Wilson, M., Zee, D. S. & Shadmehr, R. The intrinsic value of visual information affects saccade velocities. *Exp. Brain Res.* **196**, 475–481 (2009).
22. Wang, A. Y., Miura, K. & Uchida, N. The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat. Neurosci.* **16**, 639–647 (2013).
23. Mah, A., Schiereck, S. S., Bossio, V. & Constantinople, C. M. Distinct value computations support rapid sequential decisions. *Nat. Commun.* **14**, 7573 (2023).
24. Becker, J. B. Gender differences in dopaminergic function in striatum and nucleus accumbens. *Pharmacology Biochem. Behav.* **64**, 803–812 (1999).
25. Li, J. & Gibbs, R. B. Detection of estradiol in rat brain tissues: Contribution of local versus systemic production. *Psychoneuroendocrinology* **102**, 84–94 (2019).
26. Niv, Y., Joel, D. & Dayan, P. A normative perspective on motivation. *Trends Cogn. Sci.* **10**, 375–381 (2006).
27. Mah, A., Golden, C. & Constantinople, C. Mesolimbic dopamine encodes reward prediction errors independent of learning rates. *Cell Rep.* 43, 114840 (2024).
28. Starkweather, C. K., Babayan, B. M., Uchida, N. & Gershman, S. J. Dopamine reward prediction errors reflect hidden-state inference across time. *Nat. Neurosci.* **20**, 581–589 (2017).
29. Witten, I. B. et al. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72**, 721–733 (2011).
30. Millard, S. J. et al. Cognitive representations of intracranial self-stimulation of midbrain dopamine neurons depend on stimulation frequency. *Nat. Neurosci.* **27**, 1253–1259 (2024).
31. Larsen, M. B. et al. Dopamine transport by the serotonin transporter: a mechanistically distinct mode of substrate translocation. *J. Neurosci.* **31**, 6605–6615 (2011).
32. Kim, S. K. et al. Neutral sphingomyelinase 2 induces dopamine uptake through regulation of intracellular calcium. *Cell. Signal.* **22**, 865–870 (2010).
33. Watson, C. S. et al. Estradiol effects on the dopamine transporter–protein levels, subcellular location, and function. *J. Mol. Signal.* **1**, 5 (2006).
34. Fleming, W., Jewell, S., Engelhard, B., Witten, D. M. & Witten, I. B. Inferring spikes from calcium imaging in dopamine neurons. *PloS ONE* **16**, e0252345 (2021).
35. Atcherley, C. W., Wood, K. M., Parent, K. L., Hashemi, P. & Heien, M. L. The coaction of tonic and phasic dopamine dynamics. *Chem. Commun.* **51**, 2235–2238 (2015).
36. Zhang, L., Doyon, W. M., Clark, J. J., Phillips, P. E. & Dani, J. A. Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum. *Mol. Pharmacol.* **76**, 396–404 (2009).
37. Cox, J. et al. A neural substrate of sex-dependent modulation of motivation. *Nat. Neurosci.* 26, 274–284 (2023).
38. Melikian, H. E. Neurotransmitter transporter trafficking: endocytosis, recycling, and regulation. *Pharmacol. Ther.* **104**, 17–27 (2004).
39. Chen, R., Furman, C. A. & Gnegy, M. E. Dopamine transporter trafficking: rapid response on demand. *Future Neurol.* **5**, 123–134 (2010).
40. Miranda, M., Wu, C. C., Sorkina, T., Korstjens, D. R. & Sorkin, A. Enhanced ubiquitylation and accelerated degradation of the dopamine transporter mediated by protein kinase c. *J. Biol. Chem.* **280**, 35617–35624 (2005).
41. Vicentic, A., Battaglia, G., Carroll, F. I. & Kuhar, M. J. Serotonin transporter production and degradation rates: studies with RTI-76. *Brain Res.* **841**, 1–10 (1999).
42. Calipari, E. S. et al. Dopaminergic dynamics underlying sex-specific cocaine reward. *Nat. Commun.* **8**, 13877 (2017).
43. Zhang, D., Yang, S., Yang, C., Jin, G. & Zhen, X. Estrogen regulates responses of dopamine neurons in the ventral tegmental area to cocaine. *Psychopharmacol.* **199**, 625–635 (2008).
44. Shughrue, P. J., Lane, M. V. & Merchenthaler, I. Comparative distribution of estrogen receptor-α and-β mRNA in the rat central nervous system. *J. Comp. Neurol.* **388**, 507–525 (1997).
45. Creutz, L. M. & Kritzer, M. F. Estrogen receptor-β immunoreactivity in the midbrain of adult rats: regional, subregional, and cellular localization in the a10, a9, and a8 dopamine cell groups. *J. Comp. Neurol.* **446**, 288–300 (2002).
46. McHenry, J. A. et al. Hormonal gain control of a medial preoptic area social reward circuit. *Nat. Neurosci.* **20**, 449–458 (2017).
47. Bocklisch, C. et al. Cocaine disinhibits dopamine neurons by potentiation of gaba transmission in the ventral tegmental area. *Science* **341**, 1521–1525 (2013).
48. Zhu, J. & Reith, M. Role of the dopamine transporter in the action of psychostimulants, nicotine, and other drugs of abuse. *CNS Neurol. Disord. Drug Targets* **7**, 393–409 (2008).
49. Sotnikova, T. D., Beaulieu, J. -M., Gainetdinov, R. R. & Caron, M. G. Molecular biology, pharmacology and functional role of the plasma membrane dopamine transporter. *CNS Neurol. Disord. Drug Targets* **5**, 45–56 (2006).
50. Redish, A. D. Addiction as a computational process gone awry. *Science* **306**, 1944–1947 (2004).
51. Montague, P. R., Dolan, R. J., Friston, K. J. & Dayan, P. Computational psychiatry. *Trends Cogn. Sci.* **16**, 72–80 (2012).
52. Radulescu, A. & Niv, Y. State representation in mental illness. *Curr. Opin. Neurobiol.* **55**, 160–166 (2019).
53. Handy, A. B., Greenfield, S. F., Yonkers, K. A. & Payne, L. A. Psychiatric symptoms across the menstrual cycle in adult women: a comprehensive review. *Harv. Rev. Psychiatry* **30**, 100 (2022).
54. Riecher-Rössler, A., Butler, S. & Kulkarni, J. Sex and gender differences in schizophrenic psychoses-a critical review. *Arch. Womens Ment. Health* **21**, 627–648 (2018).

55. Burt, V. K., Altshuler, L. L. & Rasgon, N. Depressive symptoms in the perimenopause: prevalence, assessment, and guidelines for treatment. *Harv. Rev. Psychiatry* **6**, 121–132 (1998).

56. Kastner, P. et al. Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms a and b. *EMBO J.* **9**, 1603–1614 (1990).

57. Menegas, W., Akiti, K., Amo, R., Uchida, N. & Watabe-Uchida, M. Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nat. Neurosci.* **21**, 1421–1430 (2018).

58. Engelhard, B. et al. Specialized coding of sensory, motor and cognitive variables in vta dopamine neurons. *Nature* **570**, 509–513 (2019).

59. Takahashi, Y. K. et al. Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. *Neuron* **95**, 1395–1405 (2017).

60. Gardner, M. P., Schoenbaum, G. & Gershman, S. J. Rethinking dopamine as generalized prediction error. *Proc. Biol. Sci.* **285**, 20181645 (2018).

61. Gershman, S. J. et al. Explaining dopamine through prediction errors and beyond. *Nat. Neurosci.* **27**, 1645–1655 (2024).

62. Lee, R. S., Sagiv, Y., Engelhard, B., Witten, I. B. & Daw, N. D. A feature-specific prediction error model explains dopaminergic heterogeneity. *Nat. Neurosci.* **27**, 1574–1586 (2024).

63. Dabney, W. et al. A distributional code for value in dopamine-based reinforcement learning. *Nature* **577**, 671–675 (2020).

64. Jeong, H. et al. Mesolimbic dopamine release conveys causal associations. *Science* **378**, eabq6740 (2022).

65. Qian, L. et al. Prospective contingency explains behavior and dopamine signals during associative learning. *Nat. Neurosci.* **28**, 1280–1292 (2025).

66. Sharpe, M. J. et al. Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nat. Neurosci.* **20**, 735–742 (2017).

67. Sharpe, M. J. et al. Dopamine transients do not act as model-free prediction errors during associative learning. *Nat. Commun.* **11**, 106 (2020).

68. Jang, H. J., McMahon Ward, R., Golden, C. E. & Constantinople, C. M. Acetylcholine demixes heterogeneous dopamine signals for learning and moving. Preprint at *bioRxiv* https://doi.org/10.1101/2024.05.03.592444 (2024).

69. Kutlu, M. G. et al. Dopamine release in the nucleus accumbens core signals perceived saliency. *Curr. Biol.* **31**, 4748–4761 (2021).

70. Baek, D. -C., Kang, J. -Y., Lee, J. -S., Lee, E. -J. & Son, C. -G. Linking alterations in estrogen receptor expression to memory deficits and depressive behavior in an ovariectomy mouse model. *Sci. Rep.* **14**, 6854 (2024).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Subjects

A total of 369 Long–Evans rats (190 male, 179 female) between the ages of 6 and 24 months were used for this study (*Rattus norvegicus*), including TH-Cre ($N = 23$), ADORA2A-Cre ($N = 10$) and DRD1-Cre ($N = 3$). Animal use procedures were approved by the New York University Animal Welfare Committee (2021-1120) and carried out in accordance with National Institutes of Health (NIH) standards.

Rats were housed in pairs when possible, and water restricted to motivate them to perform behavioral trials. From Monday to Friday, they obtained water during behavioral training sessions, which were typically 90 min per day, and a subsequent ad libitum period of ~20 min. Following training on Friday until midday on Sunday, they received ad libitum water. Rats were weighed daily.

Behavioral training

We have previously published a detailed description of behavioral training for this task[23]. Reward blocks were 40 completed trials, in which the rat remained nose poking in the center port for a variable period randomly drawn from a uniform distribution from (0.8, 1.2) seconds. On each trial, an auditory tone (1, 2, 4, 8 and 16 kHz) signaled one of the five different rewards. There was a one-to-one mapping between tones and reward volumes, from smallest to largest. On each trial, the reward delay period was randomly drawn from an exponential distribution with a mean of 2.5 s; the catch probability was 15–35% until the rat matched the catch rate with opt-out rate, and then lowered to 15–25%.

Early cohorts of female rats experienced the same reward set as the males (5, 10, 20, 40 and 80 µl). However, female rats consumed less water due to their smaller body size and performed substantially fewer trials than the males. Reward offers for female rats were slightly reduced while maintaining the logarithmic spacing (4, 8, 16, 32 and 64 µl) to obtain a sufficient number of behavioral trials. For analysis, reward volumes were treated as equivalent to the corresponding volume for the male rats.

Criteria for including behavioral data. To determine when rats were sufficiently trained to understand the mapping between the auditory cues and water rewards, we evaluated their wait times on catch trials as a function of offered rewards. Only sessions with a running average slope (using a window of five sessions) above 0.01 (the average slope of shuffled data) were included to accommodate cycle-dependent changes in task performance. If fewer than three sessions per stage were included in the analysis, then these criteria were not imposed. For the comparison of water consumed over the estrous cycle (Extended Data Fig. 10), we wanted to use the same sessions for all analyses, so we imposed these criteria for behavioral performance. For the sh*Esr1* experiment, screening for behavioral performance did not provide sufficient sessions for the analysis of water consumption (Fig. 6i). Therefore, for this analysis, we included all sessions when analyzing water consumption. We emphasize that the criteria for including sessions did not evaluate rats' sensitivity to the reward blocks, or their trial initiation times.

Estrous cycle tracking. We tracked the estrous cycle for all females included in this study using vaginal cytology, with vaginal swabs collected immediately after each session using a cotton-tipped applicator first dipped in deionized water. Samples were smeared onto a clean glass slide and visually classified under bright-field illumination on a Leica ICC50 E microscope at ×10. The estrous cycle stage on each day was manually determined on the basis of the proportion of leukocytes, cornified epithelial cells and nucleated epithelial cells according to established protocols using unstained smears[71]. Given that experiments were performed over the course of months, days to weeks of irregular cycling were captured and included in the data presented.

Once the rat entered reproductive senescence, determined by a continuous estrus phase lasting longer than 2 weeks, all sessions during that time were excluded from the analyses.

To validate the accuracy of our visual classification, we measured 17β-estradiol in serum from trunk blood, collected in serum-separating tubes and spun at 3,000 rpm at 4 °C for 10 min, with the mouse/rat estradiol ELISA kit from Calbiotech (ES180S-100) after determining the stage using the method described above in 18 rats (4 in proestrus >5 h before lights out; 4 in proestrus within 3 h before lights out; 5 in estrus; 5 in diestrus), run in triplicates. Behavioral data were collected from proestrus rats during the window when their 17β-estradiol was elevated (that is, within 3 h before lights out). For this ELISA kit, the functional sensitivity (lowest concentration with a coefficient of variation (%CV) < 20%) was 3 pg ml^{-1}, the precision was 3.1% (intra-assay) and 9.9% (inter-assay), and the cross-reactivity with other endogenous steroids was 0.0002% with progesterone, 0.0001% with androstenedione, 0.0002% with testosterone and 0.0001% with cortisol[72]. All other endogenous steroids tested were <0.0001% or undetectable.

We measured serum 17β-estradiol in behaving rats using the Rat Estradiol ELISA kit from Cusabio (CSB-E05110r). For this kit, the minimum detectable amount of rat estradiol was typically less than 40 pg ml^{-1}. No significant cross-reactivity or interference between rat estradiol and analogs was observed. The intra-assay precision was CV% < 15% and the inter-assay precision was CV% < 15%. For each rat, ~200 µl blood was drawn from the lateral saphenous vein twice a week and each stage was sampled at least twice.

Behavioral model. When fitting the model to rats' initiation times (for example, to estimate RPEs), we modeled the trial initiation times as previously described (see ref. 27 for details). Briefly, we modeled initiation times as being negatively correlated to state value according to equation (1)[26,27]:

$$\text{Init}_t = D/V_t, \tag{1}$$

where V_t represents the recency-weighted average of all previous rewards and D is a scaling factor. The estimate of average reward depends on the learning rate parameter α with the recursive equation as given by equation (2):

$$V_{t+1} = V_t + \alpha\delta_t, \tag{2}$$

where V_t is the value of the state on trial t, r_t is the reward on trial t, $\delta_t = r_t - V_t$ is the RPE and α is a learning rate between 0 and 1. Because we previously found that rats exhibited a dynamic learning rate[27], we restricted model fitting to the later portion of mixed blocks, when learning rates were stable, and only related dopamine to RPEs on late mixed-block trials.

For this model, the activation function that converts values into a behavioral policy was a reciprocal function, which was chosen based on previous literature[23,26], and our finding that we could reliably recover generative parameters from this model[27]. However, when simulating this model, because the reciprocal function is nonlinear, the results were sensitive to choices of other parameters (for example, the scaling factor). Therefore, to improve the interpretability of the model simulations and make them more robust to choices about parameter values, model simulations used a linear activation function, in which the initiation times were negatively correlated with state values, parameterized with an offset and slope parameter. All model simulations used a learning rate parameter of 0.6.

Dopamine reuptake model. We modeled the activity of dopamine neurons as Poisson spikes with a background firing rate that increased or decreased (to zero) to reflect positive or negative RPEs, respectively. The dopamine concentration exhibited instantaneous increases at

action potentials, which decayed over time to reflect reuptake, according to equation (3):

$$c_t = \gamma c_{t-1} + z_t, \tag{3}$$

where c_t is the dopamine concentration, γ determines the rate of decay or reuptake and $z_t = 1$ indicates the presence of a spike at time t, or 0 otherwise. The measured dopamine fluorescence at each time point, y_t, was treated as an observation of the true concentration with Gaussian noise, according to equation (4):

$$y_t = c_t + \epsilon. \tag{4}$$

We simulated 500 trials for each condition (positive and negative RPEs), and averaged the measured dopamine signal over those trials. We performed this simulation for $\gamma = 0.65$, corresponding to faster reuptake, and $\gamma = 0.85$, corresponding to slower reuptake.

Stereotaxic surgeries. All surgeries were performed in female rats after 4 months of age with a Neurostar Robot Stereotaxic, using 3% isoflurane in oxygen at a flow rate of 2.5 l min^{-1} for induction and 2% isoflurane in oxygen at a flow rate of 1.75 l min^{-1} as maintenance for the duration of the procedure. VTA injections were conducted at a 10° angle from the midline and targeted AP −5.8; ML 0.4–0.6; DV −7.8–8.3. NAcc injections and implants targeted AP 1.3; ML 1.65; DV −6.9 with an 8–10° angle from the midline for bilateral implants.

Photometry. We used fiber photometry and the GRAB$_{DA}$ dopamine sensor (AddGene, 140554) to measure dopamine activity during the task. AAV9-hsyn-GRAB_DA2h was injected into the NAcc. To control for motion artifacts, mCherry (AAV1-CB7-CI-mCherry-WPRE-RBG, AddGene, 105-544) was also expressed because it is constitutively active. In total, 60 nl of equal parts GRAB$_{DA}$ and mCherry were delivered in each injection. Chronically implantable optic fibers (Thor labs) with a 400-µm core and 0.5-NA fiber optics were implanted unilaterally over the injection site (DV −6.7). Doric Lenses hardware and software (Doric Neuroscience Studio) were used to record fluorescence. Two-channel motion artifact correction was used to correct for movement artifacts, with mCherry as the activity-independent channel[73].

Optogenetics. We photoactivated VTA axons in the NAcc by bilaterally expressing AAV9-EF1a-double floxed-hChR2 (AddGene, 20298), 500 nl per injection, in the VTA of TH-Cre rats ($N = 5$) or simultaneously expressing AAV9-rTH-PI-Cre (AddGene, 107788, $N = 2,250$ nl ChR2 and 250 nl TH-Cre) to restrict expression to dopaminergic projections. Two tapered fibers (0.66 NA, 1-mm emitting length, 11-mm implant length, 200-µm core diameter, Optogenix) were implanted in stimulation and sham rats as controls ($N = 5$). Continuous blue light (465-nm PlexBright LEDs, Plexon) was delivered randomly on 30% of trials for 500–700 ms during the center poke, reaching 7.5 mA.

Relative protein abundance profiling. Sixteen rats (five in diestrus, five in proestrus and six in estrus), anesthetized with isoflurane, were decapitated and a brain matrix (Braintree Scientific) was used to make 1.0-mm coronal brain slices containing the central aspects of the NAcc (~0–1.6 anterior to bregma), kept on wet and dry ice. Bilateral 15-gauge punches of the NAcc were obtained and immediately placed on dry ice.

Samples were reduced with dithiothreitol at 57 °C for 1 h (5 µl of 0.2 M) and alkylated with iodoacetamide at room temperature in the dark for 45 min (5 µl of 0.5 M). Urea concentration was lowered to <2 M by addition of 100 mM ammonium bicarbonate and sequencing-grade modified trypsin (1.5 µg; Promega) was added to each sample. Digestion proceeded overnight on a shaker at room temperature. To extract proteins, samples were de-salted with Sep-Pak C18 Cartridges. C18 Cartridges were conditioned by passing 1 ml of 90% acetonitrile in

0.5% acetic acid solution three times. Then they were equilibrated by passing 3 ml of 0.1% TFA solution through them two times. Acidified samples were slowly loaded onto corresponding C18 cartridge packing. Samples were de-salted by passing 1 ml of 0.1% TFA solution through the cartridge three times. Then the cartridge was washed with 1 ml of 0.5% acetic acid. Peptides were eluted into clean tubes by slowly passing 500 µl of 40% acetonitrile in 0.5% acetic acid, followed by the addition of 80% acetonitrile in 0.5% acetic acid.

Samples were reconstituted in 30 µl of 100 mM HEPES solution, pH ~8.5, and labeled with TMT dissolved in 90 µl of 100% anhydrous ethanol for 5 min. Each sample was combined with its corresponding TMT label and incubated for 1 h. The reaction was quenched by adding 50 µl of 100 mM ammonium bicarbonate solution and incubating for 15 min. Samples were dried in a SpeedVac concentrator. The sample was then fractionated into 45 fractions using basic high-performance liquid chromatography fractionation, and the fractions were concatenated into 15 final fractions for liquid chromatography coupled with mass spectrometry (LC–MS) analysis where 1/10th of the fractions were analyzed individually. Liquid chromatography separation online with a mass spectrometer was performed using the autosampler of an EASY-nLC 1000 (Thermo Scientific). Peptides were gradient eluted from the column directly to an Eclipse mass spectrometer using a 1-h gradient (Thermo Scientific) with 2% acetonitrile and 0.5% acetic acid for solvent A and 80% acetonitrile and 0.5% acetic acid for solvent B.

High-resolution full mass spectrometry spectra were acquired with a resolution of 60,000, an automatic gain control target of 4×10^5, with a maximum ion time of 50 ms, and scan range of 400 to 1,500 m/z. Following each full mass spectrometry scan, the top ten data-dependent higher-energy collisional dissociation tandem mass spectrometry (MS/MS) spectra were acquired. All MS/MS spectra were collected using the following instrument parameters: resolution of 60,000, automatic gain control target of 1×10^5, maximum ion time of 60 ms, one microscan, 0.7-m/z isolation window and normalized collision energy of 30. MS/MS spectra were searched against the rat database using MaxQuant.

Only proteins that had at least two peptides detected and that were identified in at least three samples of one estrous stage were included in the analysis. In total, 3,082 proteins were identified and 3,043 were quantified. Reporter ion intensity values were normalized and all protein intensities were log$_2$ transformed. The mass spectrometric raw files are accessible at https://massive.ucsd.edu/ under accession MassIVE MSV00098185.

Electron microscopy. Rats ($N = 12$) were anesthetized with isoflurane and perfused through the ascending aorta with 50 ml of phosphate buffered saline (PBS) containing 0.2 ml of heparin and then 4% paraformaldehyde dissolved in 0.1 M phosphate buffer, pH 7.4. The brains were removed from the calvariae and post-fixed for one week in 4% paraformaldehyde. Sections through the NAcc (50 µm) were cut with a vibratome, and stored in 0.01 M PBS and 0.05% azide. To enhance the antibody's penetration of tissue, free-floating vibratome sections underwent a freeze/thaw protocol, for which tissue was cryo-protected in increasing concentrations of dimethyl sulfoxide (DMSO; 5%, 10% and 20%) for 10 min each before rapid freezing in a beaker of 4-methylbutane chilled using dry ice and 100% ethanol as the heat conduit. After rapid freezing, the tissue was thawed by immersing in a bath of 20% DMSO at room temperature, before another bout of rapid freezing for a total of eight freeze/thaw sequences. Tissue was then incubated in decreasing concentrations of DMSO (10% and 5%) for 5 min each before returning to PBS[74]. Tissue was stained for DAT using the protocol described below, except with an anti-DAT concentration of 1:250 dilution that the tissue was treated with for 48 h. Secondary antibody incubation was overnight in PBS-BSA-azide containing ultra-small colloidal gold-conjugated goat anti-rabbit secondary antibody (1:100 dilution; EMSciences, 25181, lot no. GAR-01217/1). Sections were rinsed

in PBS and post-fixed in 2% glutaraldehyde in PBS for 10 min, and then the steps for silver intensification (KPL Silver Enhancer Kit, 5520-0021) were carried out to enlarge the ultra-small colloidal gold particles.

The electron microscopy tissue processing was similar to the procedure performed before[75], which consisted of osmium-free tissue processing in 1% tannic acid, 1% uranyl acetate and 0.2% iridium tetrabromide, each dissolved in maleate buffer (pH 6.0, 0.1 M), post-fixation in 1% uranyl acetate in 70% ethanol overnight, dehydration in 70%, 90% and 100% ethanol, followed by 100% acetone, infiltration in EPON 812 (EMSciences), flat-embed between Aclar plastic sheets (EMSciences) and capsule-embedded in BEEM capsules (EMSciences). Ultrathin sections were prepared using the Ultracut E ultramicrotome and collected on nickel grids (EMSciences). Sections were counter-stained with lead citrate to improve contrast. Images were captured at a magnification of ×50,000 using the Hamamatsu CCD camera (1.2 megapixels) and software from AMT.

For each animal, approximately 50 axons were sampled and quantified. Axons were considered to be DAT positive when one or more SIG particles could be detected along its plasma membrane or within its cytosol. Axons were identified to be forming symmetrical or asymmetrical synapses. SIG particles were categorized as either membranous (center of particle was perfectly aligned to plasma membrane) or intracellular (center of particle was within the boundaries of the axon). The size of the axons and SIG particles were assessed by measuring the maximum and minimum diameters. The distance of each SIG particle to the plasma membrane was measured using the center of the SIG particle and the nearest point along the membrane. A total of 897 SIG particles were included in the analysis.

Estrogen receptor suppression. AAV2-U6/TO-sh*Esr1*-CMV-TetR-P2A-eGFP-KASH-pA was custom constructed from AAV backbone (AddGene, 82706)[76] with the insertion of the shRNA sequence of *Esr1* (ref. [77]) and custom prepped by Wz Biosciences. pJEP11-AAV-U6/TO-sgRNA(Tet2)-CMV-TetR-P2A-eGFP-KASH-pA was a gift from J. Ploski (AddGene plasmid no. 82705). sh*Esr1* was also cloned into the lentiviral vector pLL3.7 as previously described[78,79] (AddGene plasmid, 120720). Either AAV-U6TO-sh*Esr1*-CMV-TetR-eGFP (U6/tetO-sh*Esr1*, $N = 5$) or pLL3.7-sh*Esr1* (lenti-sh*Esr1*, $N = 4$), which use the same sh*Esr1* sequence[77], was injected bilaterally into the VTA in eight rats to knock down *Esr1* expression. Dox was used to induce viral expression of U6/tetO-sh*Esr1*, delivered via both ad lib food (Bio-Serv diet) and 10–20 min of ad lib dox-containing water (~3 mg ml⁻¹; Sigma-Aldrich dox powder) such that each rat received 20 mg per kg body weight per day (below toxic levels)[80]. All behavioral sessions after the first day of the dox regimen were included in the analysis for U6/tetO-sh*Esr1*-injected rats and after 21 days of expression for lenti-sh*Esr1*-injected rats and compared to sessions that preceded surgery.

Serum osmolality. Osmolality was measured in the same serum samples used to detect 17β-estradiol expression. The extracted serum was sent by refrigerated transportation to a commercial laboratory for osmolality determination using a freezing-point method (Cornell University Clinical Pathology Laboratory).

Histological verification with protein and RNA staining. Rats were euthanized following completion of behavioral testing for photometry, optogenetics and shRNA experiments and when they reached adulthood (10–11 weeks) for electron microscopy and immunohistochemistry quantification experiments. Rats were first perfused with PBS and subsequently with 4% paraformaldehyde for DAT and SERT quantification, and 10% formalin for histological verification for optogenetics, photometry and sh*Esr1* experiments. Brains were post-fixed with 4% paraformaldehyde for one week at 4 °C for DAT and SERT quantification, with 10% formalin for 2 days for verification of surgical coordinates, and with 10% formalin for one day for sh*Esr1*

validation. Sections (50 μm) of extracted brain tissue were sliced on a Leica VT1000 vibratome.

Sections were treated with 1% hydrogen peroxide in 0.01 M PBS for 30 min, washed with 0.01 M PBS and blocked in 0.01 M PBS containing 0.05% sodium azide and 1% bovine serum albumin (PBS-BSA-azide) for 30 min before being treated with primary antibody overnight at room temperature. Primary antibody was removed with PBS rinses and then sections were treated with secondary antibodies for 1 h away from light. Primary and secondary antibodies were made in PBS-BSA-azide and primary antibodies additionally contained 0.2% Triton-X. To verify virus expression and implant location, endogenous fluorescent markers were amplified with rabbit anti-GFP (Thermo Fisher Scientific A11122, 1:2,000 dilution, lot no. 2083201) and imaged with an Olympus VS120 Virtual Slide Microscope. For light microscopy experiments, DAT was labeled with a 1:500 dilution of Sigma-Aldrich D6944 antibody (lot no. 0000124874), SERT was labeled with a 1:2,000 dilution of Chemicon AB1594P antibody (lot no. 18020403), and TH was labeled with a 1:400 dilution of Millipore MAB318 antibody (lot no. 3990619). The secondary antibodies used were as follows (all 1:200 dilution and from Thermo Fisher Scientific): anti-rabbit Alexa Fluor 488 IgG A11008 (lot no. 2147635) for DAT, SERT and GFP, and goat anti-mouse Alexa Fluor 555 for TH (A21424, lot no. 2390715). All slices were mounted on glass slides using Prolong Gold (Vector Laboratories).

For verification of *Esr1* knockdown by U6/tetO-sh*Esr1* and lenti-sh*Esr1* viruses, RNA FISH was performed using the ACD Bio RNAscope Multiplex Fluorescent v2 assay. Brains were cryoprotected in a solution of 10% sucrose/PBS and stored at 4 °C until the brain sank. Cryoprotection was repeated with 20% and 30% sucrose/PBS solutions. The brains were embedded in optimal cutting temperature solution and flash frozen with 2-methylbutane cooled to −80 °C. For tissue sectioning, a cryostat was used to collect 10-μm sections stamped onto superfrost plus slides. The sections were air-dried at −20 °C for 1 h and stored at −80 °C. Sections were stained with target probes EGFP-O4-C1 (ACD Bio 538851, lot no. 24117A) to label GFP and Rn-Esr1-C2 (ACD Bio 317151-C2, lot no. 24143D) diluted to 1:50 dilution to label *Esr1*. RNAscope Multiplex Fluorescent Reagent Kit v2 (ACD Bio 323270) was used for the FISH assay. Control tissue was labeled with 3-Plex Positive Control Probe (ACD Bio 320891, lot no. 24121A) and 3-Plex Negative Control Probe (ACD Bio, 320871, lot no. 2021233) as internal controls. DAPI was used as a counterstain.

Quantification of fluorescence. Photomicrographs were taken at a magnification of ×20 and ×63 with a Leica DM6 CS Confocal using standardized parameters developed to maximize signal-to-noise ratios. z-stacks were taken for *Esr1* quantification at a step size of 0.3 μm.

Photomicrographs were analyzed in ImageJ to quantify protein expression. A mask was generated to identify the location of fluorescing pixels. The total area of the mask (μm²) was used to determine expression levels of DAT and SERT.

FISH-quant was used for *Esr1* quantification[81]. Detection settings were set based on filtering of background noise and accurate identification of spots on internal control sections, and standardized across all sections. The total number of spots in each image of a z-stack was used for analysis.

Statistics and reproducibility
General. No statistical method was used to predetermine sample size. Effect sizes (Cohen's *d*) for effects across the estrous cycle, which had matching sample sizes across groups, were calculated as given by equation (5):

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}} \tag{5}$$

For independent samples, Cohen's *d* was calculated as given by equation (6):

$$d = \frac{\mu_1 - \mu_2}{s_1} \qquad (6)$$

Trial initiation times. Trial initiation time was defined as the duration of time in between the termination of one trial (violation, opt out, or the last lick from the rewarded port) and the subsequent center poke that began the next trial. We excluded outlier trial initiation times that were above the 98th percentile of the rat's cumulative trial initiation time distribution pooled over sessions, except when less than three sessions per stage were used, then the cutoff was the 99th percentile. Sensitivity of trial initiation times to blocks was defined as the change in initiation time (mean in low−mean in high blocks). Where noted, the effect of satiety over the session on trial initiation times was accounted for by subtracting the regression of mean trial initiation time as a function of trial number.

Trial history effects. To assess trial initiation time sensitivity to previous offers, we regressed trial initiation time against previous rewards. We focused only on mixed blocks. Reward offers from a different block (for example, a previous high block) were given not a number (NaN) values. Then, we regressed against the previous nine reward offers, not including the current trial, along with a constant offset. We set the reward for violation, catch and opt-out trials to zero, because rats do not receive a reward on these trials. We used MATLAB's built-in regress function to perform the regression. With the coefficients, we found the first nonsignificant coefficient (coefficient with a 95% confidence interval containing zero), and set that coefficient and all following coefficients to zero, as these coefficients cannot meaningfully be interpreted as being different from zero. To determine the relationship between the estrous cycle and the effect of trial history on initiation times, we created a multiple linear regression model that included cycle stage in addition to reward history. When only two stages were considered, diestrus was set to 1 and proestrus was set to 2. When all four stages were considered, diestrus was set to 1, metestrus was set to 2, estrus was set to 3, and proestrus was set to 4 to parallel the cycle of 17β-estradiol and the distance from its peak.

Wait times. Wait times on catch trials were included in the analysis. Wait times greater than three times the standard deviation above the mean were considered outliers and excluded.

Photometry. Trials that followed violations were excluded from the data presented in Fig. 2.

Optogenetics. Only control sessions that occurred before the first day of optogenetic stimulation were included in the analysis because optogenetic activation of dopamine terminals had effects on behavior that lasted beyond the day of stimulation. Initiation times were normalized to high blocks in control sessions for both optogenetic and sham groups.

Proteomics. Proteins were determined to be significantly differentially expressed with a *t*-test and a Benjamini–Hochberg false discovery rate of 5% to correct for multiple comparisons. Significantly differentially expressed protein signatures identified with mass spectrometry were subjected to functional annotation. STRING database v12 was used to assess enrichment of Gene Ontology terms for *Rattus norvegicus* using a hypergeometric test and all identified proteins as background. According to https://geneontology.org/, 246 genes and their proteins are associated with dopamine; 92 of these proteins were identified by mass spectrometry, and 3 that were significantly decreased in expression are related to dopamine uptake and transport (Slc6a3, Slc6a4, and Smpd3).

Serum osmolality. Osmolality measurements greater than 2.5 times the standard deviation above the mean were considered outliers and excluded. 17β-Estradiol expression for one rat in estrus was excluded when correlated with osmolality due to it being an outlier (less than one standard deviation below the mean of all samples).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The mass spectrometric raw files are accessible at https://massive.ucsd.edu/ under accession MassIVE MSV00098185. The rest of the data used in this study are deposited in Zenodo via https://doi.org/10.5281/zenodo.16903822 (ref. 82) (Figs. 1, 5 and 6), https://doi.org/10.5281/zenodo.16912311 (ref. 83) (Fig. 2), https://doi.org/10.5281/zenodo.16912773 (ref. 84) (Fig. 3), https://doi.org/10.5281/zenodo.16912863 (ref. 85) (Fig. 4) and https://doi.org/10.5281/zenodo.16914371 (ref. 86), https://doi.org/10.5281/zenodo.16913033 (ref. 87) and https://doi.org/10.5281/zenodo.16914534 (ref. 88) were used for Extended Data.

Code availability

Code used to analyze the data and generate the figures is available at https://github.com/constantinoplelab/published/tree/main/EstrousRPEPaper/.

References

71. Marcondes, F., Bianchi, F. & Tanno, A. Determination of the estrous cycle phases of rats: some helpful considerations. *Brazilian J. Biol.* **62**, 609–614 (2002).

72. Haisenleder, D. J., Schoenfelder, A. H., Marcinko, E. S., Geddis, L. M. & Marshall, J. C. Estimation of estradiol in mouse serum samples: evaluation of commercial estradiol immunoassays. *Endocrinology* **152**, 4443–4447 (2011).

73. Creamer, M. S., Chen, K. S., Leifer, A. M. & Pillow, J. W. Correcting motion induced fluorescence artifacts in two-channel neural imaging. *PLoS Comput. Biol.* **18**, e1010421 (2022).

74. Santiago, A. N., Lim, K. Y., Opendak, M., Sullivan, R. M. & Aoki, C. Early life trauma increases threat response of peri-weaning rats, reduction of axo-somatic synapses formed by parvalbumin cells and perineuronal net in the basolateral nucleus of amygdala. *J. Comp. Neurol.* **526**, 2647–2664 (2018).

75. Chen, Y. -W., Wable, G. S., Chowdhury, T. G. & Aoki, C. Enlargement of axo-somatic contacts formed by gad-immunoreactive axon terminals onto layer V pyramidal neurons in the medial prefrontal cortex of adolescent female mice is associated with suppression of food restriction-evoked hyperactivity and resilience to activity-based anorexia. *Cerebral Cortex* **26**, 2574–2589 (2016).

76. De Solis, C. A., Ho, A., Holehonnur, R. & Ploski, J. E. The development of a viral mediated CRISPR/Cas9 system with doxycycline dependent grna expression for inducible in vitro and in vivo genome editing. *Front. Mol. Neurosci.* **9**, 70 (2016).

77. Musatov, S., Chen, W., Pfaff, D. W., Kaplitt, M. G. & Ogawa, S. Rnai-mediated silencing of estrogen receptor *α* in the ventromedial nucleus of hypothalamus abolishes female sexual behaviors. *Proc. Natl Acad. Sci. USA* **103**, 10456–10460 (2006).

78. Lasek, A., Janak, P., He, L., Whistler, J. & Heberlein, U. Downregulation of mu opioid receptor by RNA interference in the ventral tegmental area reduces ethanol consumption in mice. *Genes Brain Behav.* **6**, 728–735 (2007).

79. Satta, R., Certa, B., He, D. & Lasek, A. W. Estrogen receptor *β* in the nucleus accumbens regulates the rewarding properties of cocaine in female mice. *Int. J. Neuropsychopharmacol.* **21**, 382–392 (2018).

80. El-Neweshy, M. S. Experimental doxycycline overdose in rats causes cardiomyopathy. *Int. J. Exp. Pathol.* **94**, 109–114 (2013).

81. Mueller, F. et al. Fish-quant: automatic counting of transcripts in 3D fish images. *Nat. Methods* **10**, 277–278 (2013).

82. Golden, C. E. M. et al. Estrogen modulates reward prediction errors and reinforcement learning. *Zenodo* https://doi.org/10.5281/zenodo.16903822 (2025).

83. Golden, C. E. M. et al. Estrogen modulates reward prediction errors and reinforcement learning. *Zenodo* https://doi.org/10.5281/zenodo.16912311 (2025).

84. Golden, C. E. M. et al. Estrogen modulates reward prediction errors and reinforcement learning. *Zenodo* https://doi.org/10.5281/zenodo.16912773 (2025).

85. Golden, C. E. M. et al. Estrogen modulates reward prediction errors and reinforcement learning. *Zenodo* https://doi.org/10.5281/zenodo.16912863 (2025).

86. Golden, C. E. M. et al. Estrogen modulates reward prediction errors and reinforcement learning. *Zenodo* https://doi.org/10.5281/zenodo.16914371 (2025).

87. Golden, C. E. M. et al. Estrogen modulates reward prediction errors and reinforcement learning. *Zenodo* https://doi.org/10.5281/zenodo.16913033 (2025).

88. Golden, C. E. M. et al. Estrogen modulates reward prediction errors and reinforcement learning. *Zenodo* https://doi.org/10.5281/zenodo.16914534 (2025).

Author contributions

C.M.C. and C.E.M.G. designed the study. C.E.M.G., A.C.M., D.K. and D.H.L. collected and analyzed the data. A.M. developed the behavioral model and contributed to data analysis. T.Y. and D.L. created and shared the U6/tetO-sh*Esr1* virus. A.W.L. created and shared the lenti-sh*Esr1* virus. C.A. supervised the electron microscopy. C.E.M.G. prepared the figures. C.E.M.G. and C.M.C. wrote the paper. C.M.C. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41593-025-02104-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-025-02104-z.

Correspondence and requests for materials should be addressed to Christine M. Constantinople.

Peer review information *Nature Neuroscience* thanks Melissa Sharpe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.
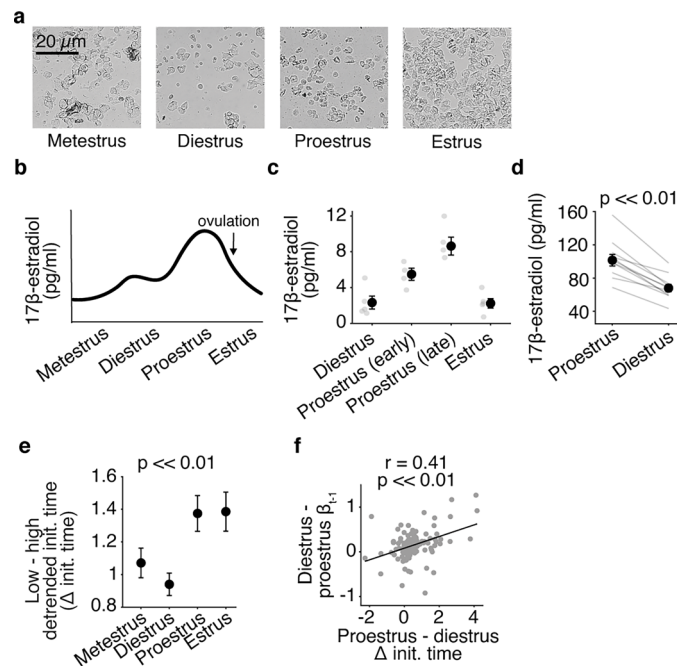
Extended Data Fig. 1 | Trial initiation times are most modulated by the value of the state around the time of ovulation. a. Example images of vaginal cytology by estrous stage. **b-c.** Expected **(b)** and observed **(c)** median 17β-estradiol, measured with an ELISA from serum, $N = 18$ rats (4 in proestrus > 5 hours before lights out; 4 in proestrus within 3 hours before lights out; 5 in estrus; 5 in diestrus), Kruskal-Wallis test $p = 6.15 \times 10^{-3}$ for group effect, two-sided Wilcoxon rank sum test for late proestrus vs. diestrus $p = 0.016$ and $d = 3.74$, early proestrus vs. late proestrus $p = 0.029$, estrus vs. early proestrus $p = 0.032$, estrus vs. late proestrus $p = 0.016$, diestrus vs. estrus $p = 1$, diestrus vs. early proestrus $p = 0.063$. **d.** Serum 17β-estradiol is significantly higher in proestrus than diestrus during training sessions, $N = 17$ rats, two-sided Wilcoxon signed rank test $p = 9.77 \times 10^{-4}$,

$d = 1.79$. **e.** Median sensitivity of detrended trial initiation times to blocks for each estrous stage, $N = 120$ female rats, Kruskal-Wallis test $p = 5.00 \times 10^{-5}$ for group effect, two-sided Wilcoxon signed-rank tests for post-hoc analyses: $p = 0.002$ and $d = 0.17$ for proestrus vs. metestrus, $p = 2.67 \times 10^{-8}$ and $d = 0.41$ for proestrus vs. diestrus, $p = 0.09$ and $d = 0.17$ for proestrus vs. estrus, $p = 9.06 \times 10^{-11}$ and $d = 0.32$ for estrus vs. metestrus, $p = 2.67 \times 10^{-15}$ and $d = 0.53$ for estrus vs. diestrus, and $p = 9.34 \times 10^{-4}$ and $d = 0.20$ for metestrus vs diestrus). **f.** The differences in regression coefficients of one trial back for proestrus and diestrus are significantly correlated with the differences in block sensitivity for proestrus and diestrus, two-sided Pearson correlation $p = 4.17 \times 10^{-6}$. All circles are medians and error bars are median ± SEM.
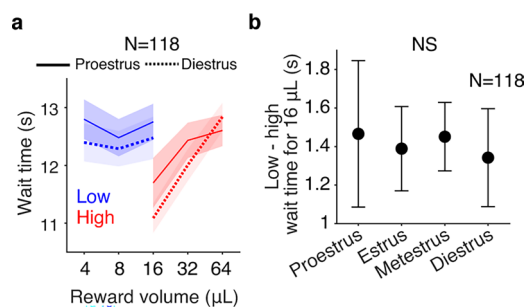
Extended Data Fig. 2 | Wait time adaptation to reward context does not change across the cycle. a. Mean wait time on catch trials by reward in each block (blue = low, red = high) averaged across rats. Error bars are mean ± SEM. **b**. Median adaptation of wait times to reward blocks (low - high block) does not change across the cycle, $N = 118$ rats, Kruskal-Wallis test $p = 0.69$ for group effect and two-sided Wilcoxon signed-rank test $p = 0.771$ and $d = 0.06$ for proestrus vs. diestrus. Error bars are median ± SEM.
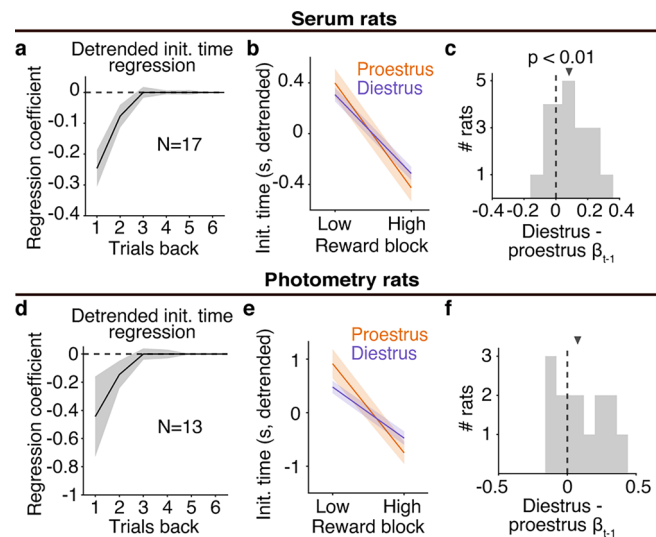
Extended Data Fig. 3 | Population-level effects of the estrous cycle hold in rats used in serum 17β-estradiol and photometry experiments. a-c. Rats used in serum 17β-estradiol experiments have (**a**) median regression coefficients of detrended trial initiation times plotted against rewards on previous trials during mixed blocks that decline as an exponential function of time, (**b**) enhanced block sensitivity in proestrus, and (**c**) greater regression coefficients for the previous reward in proestrus because the distribution is significantly different from zero using a two-sided Wilcoxon signed-rank test $p = 4.88 \times 10^{-3}$. **d-f.** Rats used in photometry experiments have (**d**) median regression coefficients of detrended trial initiation times plotted against rewards on previous trials during mixed blocks that decline as an exponential function of time, (**e**) enhanced block sensitivity in proestrus, and (**f**) greater regression coefficients for the previous reward in proestrus. All arrows are medians and all error bars are median ± SEM.
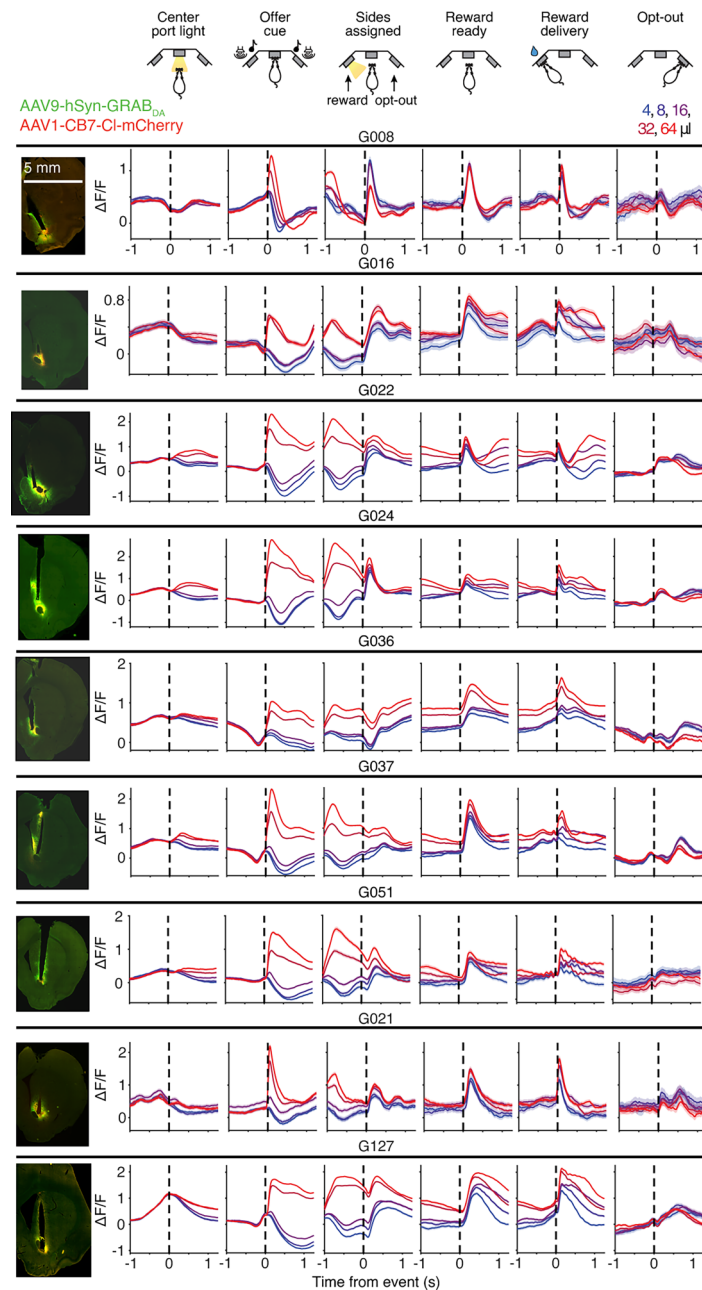
Extended Data Fig. 4 | Histology and task-aligned dopamine across thirteen individual rats. For each rat, histological verification of fiber implant targeting and mCherry (red) and ChR2 (green) expression (left) and task event-aligned dopamine responses, split by reward offer cue volume (right). Data is not baseline-corrected. All error bars are mean ± SEM.
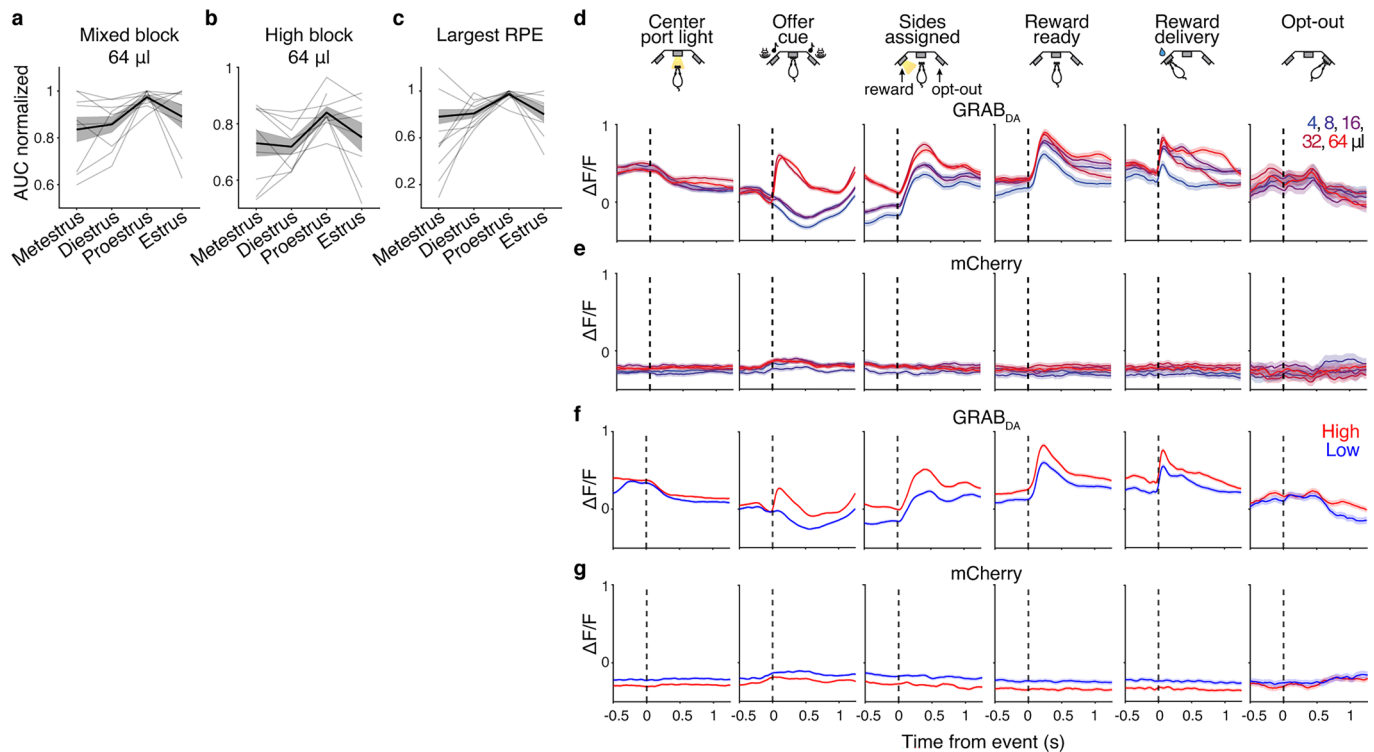
Extended Data Fig. 5 | RPE is most strongly encoded in proestrus compared to the rest of the estrous cycle and not encoded by mCherry. Normalized AUC of mean dopamine response across the cycle to largest reward offer cue in mixed blocks (**a**), Kruskal-Wallis test $p = 0.048$ for group effect and two-sided Wilcoxon signed-rank test $p = 3.91 \times 10^{-3}$ for proestrus vs. diestrus, (the rest of the pairwise comparisons were not significant), and high blocks (**b**), Kruskal-Wallis test $p = 0.15$ for group effect, and largest model-predicted RPE bin (**c**), Kruskal-Wallis

test $p = 0.015$ for group effect and two-sided Wilcoxon signed-rank test $p = 0.039$ for metestrus vs. proestrus, $p = 0.027$ for diestrus vs. proestrus, and $p = 0.023$ for proestrus vs. estrus (the rest of the pairwise comparisons were not significant). **d-e**. While motion-corrected GRAB$_{DA}$ encodes reward volume, mCherry, used to control for motion artifacts, does not. **f-g**. Same for block encoding. All error bars are mean ± SEM.
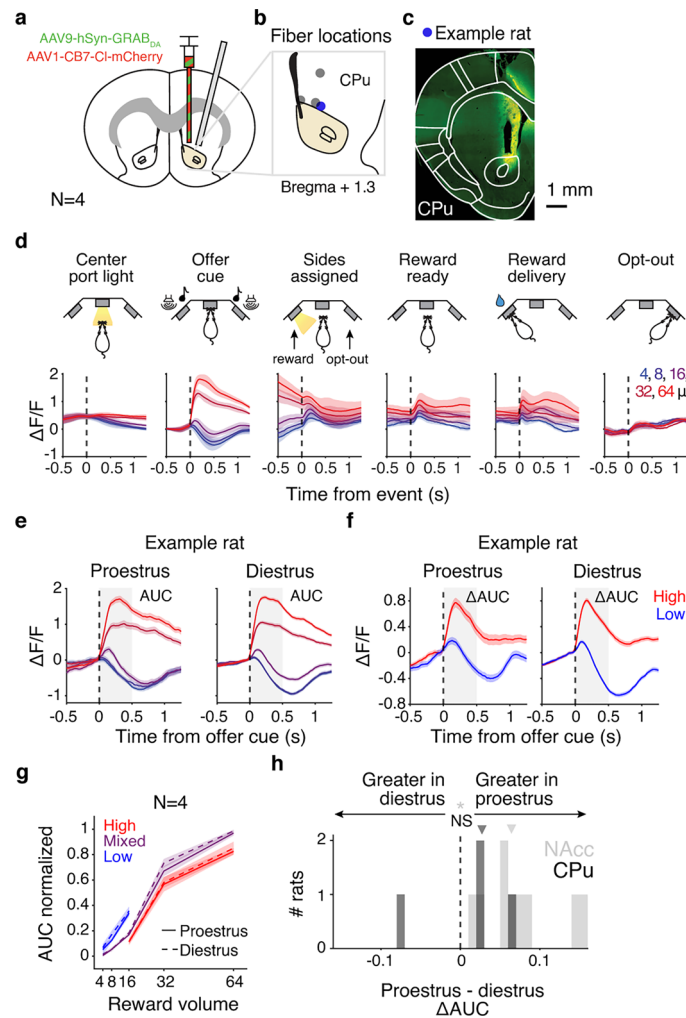
Extended Data Fig. 6 | Dopamine in audate putamen above NAcc encodes RPE, but is unaffected by estrous. a-b. Fiber locations for 4 rats. Blue circle in **b** is example depicted in **c**. CPu = caudate putamen. **c**. Example histology of blue circle in **b**. **d**. Task event-aligned dopamine signaling, separated by reward volume. **e-f**. Response to offer cue, separated by reward volume **(e)** and block **(f)** and stage group for example rat, G027. Gray boxes represent the window used to calculate change in area under curve, 0 to 0.5 s (ΔAUC) in **(g)** and **(h)**. Data is baseline-corrected using the 0.05 to 0 s before offer cue. **g**. No significant difference in median population response to each reward volume, separated by reward block, min-max normalized, and baseline-corrected using the 50 ms before offer cue. **h**. Histogram of stage effect (proestrus-diestrus) on ΔAUC during offer cue for all NAcc and CPu rats (*p = 3.9 × 10^{-3} for NAcc using a two-sided Wilcoxon signed-rank test, p = 0.875 for CPu). Triangles are medians. All error bars are mean ± SEM.
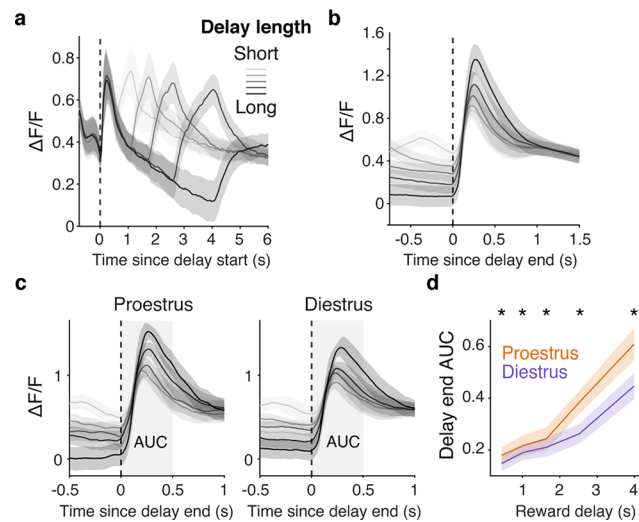
Extended Data Fig. 7 | Delay end RPE is enhanced in proestrus. a. Mean dopamine response aligned to beginning of delay period for rewarded mixed block trials by reward delay length. **b**. Mean dopamine response aligned to the end of the delay period for rewarded mixed block trials by reward delay length. **c**. Dopamine 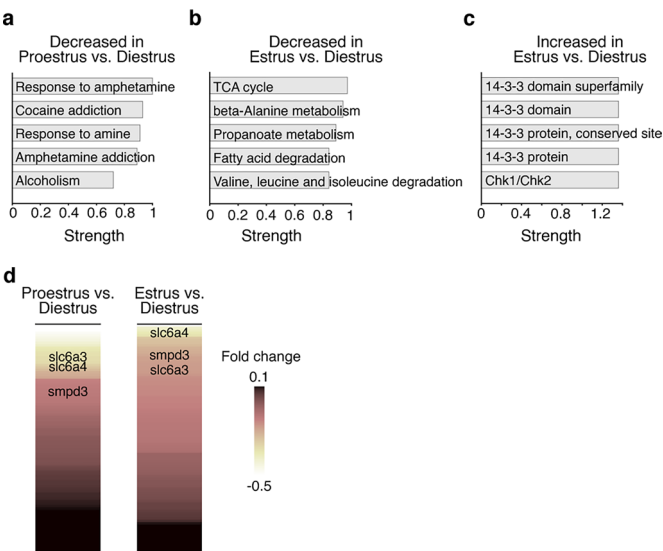response aligned to the end of the delay period when reward was cued for mixed block trials by reward delay length and stage. **d**. Median AUC from c is greater in proestrus than diestrus, two-sided Wilcoxon signed-rank test (from low to high delay) $p = 7.80 \times 10^{-3}$, 0.02, 0.01, $p = 7.80 \times 10^{-3}$, $p = 3.90 \times 10^{-3}$. Error bars are ± SEM.
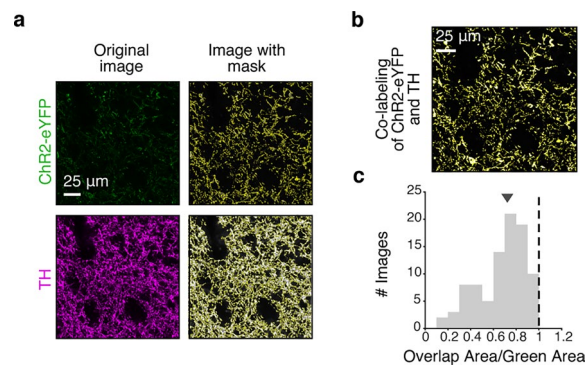
Extended Data Fig. 8 | Expression of ChR2 in TH+ axons terminals in the NAcc.
a. Example images from immunohistochemical quantification of ChR2-eYFP (top) and TH (bottom) co-labeling at 63x with mask (right) identifying fluorescent pixels above a threshold, generated in ImageJ, overlaid. Performed on 90 regions of interest. **b**. Image created in ImageJ that depicts pixels that are fluorescent in both ChR2-eYFP and TH images (overlap) with mask. **c**. Metric of co-labeling across all images: area of mask segmentation for ChR2-eYFP image divided by area of mask segmentation for TH image. The arrow indicates the median.

a

Decreased in
Proestrus vs. Diestrus

Response to amphetamine
Cocaine addiction
Response to amine
Amphetamine addiction
Alcoholism

0   0.2 0.4 0.6 0.8   1
Strength

b

Decreased in
Estrus vs. Diestrus

TCA cycle
beta-Alanine metabolism
Propanoate metabolism
Fatty acid degradation
Valine, leucine and isoleucine degradation

0   0.2 0.4 0.6 0.8   1
Strength

c

Increased in
Estrus vs. Diestrus

14-3-3 domain superfamily
14-3-3 domain
14-3-3 protein, conserved site
14-3-3 protein
Chk1/Chk2

0   0.4   0.8   1.2
Strength

d

Proestrus vs. Diestrus

Estrus vs. Diestrus

Fold change
0.1

-0.5

Extended Data Fig. 9 | Proteins related to dopamine reuptake are significantly decreased in proestrus and estrus. a-c. The strongest gene ontology terms that decreased **(a-b)** and increased **(c)** proteins were significantly enriched for in proestrus **(a)** and estrus **(b-c)** compared to diestrus. Increased proteins in proestrus compared to diestrus were not significantly enriched for any known biological functions. **d.** 92 differentially expressed proteins related to dopamine sorted by fold change, with proteins involved in dopamine reuptake identified.
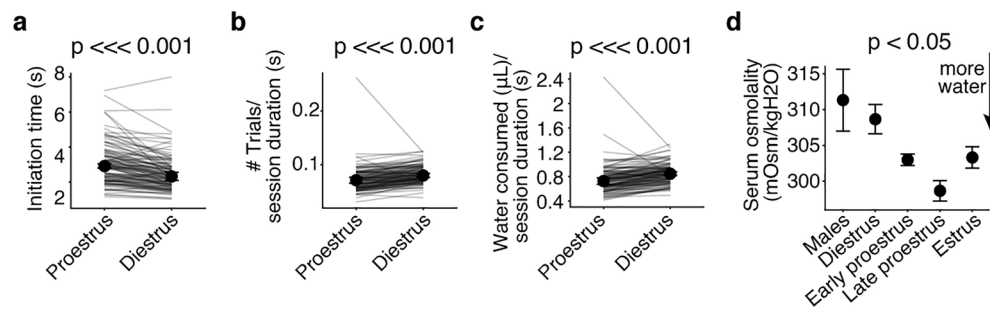
Extended Data Fig. 10 | Systemic expression of 17β-estradiol correlates with reduced thirst. a. Median initiation time across all trials is significantly higher in proestrus, $N = 118$ rats, two-sided Wilcoxon signed rank test $p = 1.60 \times 10^{-14}$ and $d = 0.49$. **b-c.** Median trials performed (**b**), $p = 8.30 \times 10^{-16}$ and $d = 0.48$, and total water consumed (**c**), $p = 1.34 \times 10^{-15}$ and $d = 0.51$, controlling for the total amount of time the rat was allotted in the behavioral rig, $N = 118$ rats. **d.** Median serum osmolality was lower for stages around ovulation (proestrus and estrus, $N = 5$ males, $N = 7$ diestrus, $N = 4$ early proestrus, $N = 4$ late proestrus, and $N = 5$ estrus), Kruskal-Wallis test $p = 0.01$ for the group effect and $d = 1.61$ for proestrus vs diestrus. All error bars are median $\pm$ SEM.

Corresponding author(s): Christine Constantinople

Last updated by author(s): 6/29/2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Code used to collect behavioral data with Bpod was created in Matlab and is available at https://github.com/constantinoplelab/published/tree/main/EstrousRPEPaper.
Data analysis	Code used to analyze data and generate figures was created in Matlab and is available at https://github.com/constantinoplelab/published/tree/main/EstrousRPEPaper.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The mass spectrometric raw files are accessible at https://massive.ucsd.edu under accession MassIVE MSV00098185. The rest of the data used in this study were

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We did not perform a calculation on sample size. We believe our sample size is adequate because it is comparable to or larger than many similar studies
Data exclusions	Performance criteria was used to select sessions for behavioral analyses in order to exclude sessions where rats were still learning the task; behavioral sessions were also excluded when female rats were no longer cycling; outlier trial initiation times that were above the 98th percentile of the rat's cumulative trial initiation time distribution pooled over sessions were excluded, or 99th for 3 sessions or less per rat; outlier delta trial initiation times more than 1.5 x the standard deviation were excluded for the correlation with estradiol; measurements of serum osmolality greater than 2.5 x the standard deviation were excluded (two samples); estradiol expression for one rat in estrus was excluded when correlated with osmolality due to it being one standard deviation below the mean of all samples.
Replication	Data were collected in multiple, sequential cohorts of males and females that spanned four years and results were consistent across cohorts. Imaging data was collected separately, as the animals completed the experiments, but was analyzed in a single cohort to standardize measurements. We did not take cohort into account for statistical analyses.
Randomization	Rats were randomly assigned to experimental groups (ChR2, sham, doxycycline, no doxcycline).
Blinding	Experimenters could not be blinded to sex during data collection because rats are housed by sex; all behavioral and imaging data was recorded and measured by automated hardware and software so that the experimenter had no effect on the outcome.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☐ Eukaryotic cell lines
☐	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☐ Clinical data
☐	☐ Dual use research of concern
☐	☐ Plants

Methods

n/a	Involved in the study
☐	☐ ChIP-seq
☐	☐ Flow cytometry
☐	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For light microscopy experiments, DAT was labeled with 1:500 Sigma-Aldrich D6944 lot #0000124874, SERT was labeled with 1:2,000 Chemicon AB1594P lot #18020403, TH was labeled with 1:400 Millipore MAB318 lot #3990619, and ER alpha was labeled with 1:250 Thermo Fisher Scientific PA1-309 lot #YG378288. The secondary antibodies used were as follows (all 1:200 and from Thermo Fisher Scientific): anti-rabbit AlexaFluor 488 IgG (A11008, lot #2147635) for DAT, SERT, and GFP, goat anti-mouse AlexaFluor 555 for TH (A21424, lot #2390715), and goat anti-rabbit AlexaFluor 594 (A11037, lot #2160431) for ER alpha.
Validation	All antibodies were validated by the respective manufacturers to work for immunohistochemistry, as stated on their online product pages.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.
Authentication	Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.
Mycoplasma contamination	Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology and Archaeology

Specimen provenance	Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.
Specimen deposition	Indicate where the specimens have been deposited to permit free access by other researchers.
Dating methods	If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	A total of 366 Long-evans rattus norvegus rats (201 male, 165 female) between the ages of 6 and 24 months were used for this study, including ADORA2A-Cre (N=10), DRD1-Cre (N=3), and TH-Cre (N=23).

Wild animals	The study did not involve wild animals.
Reporting on sex	All results are reported by sex.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	Animal use procedures were approved by the New York University Animal Welfare Committee (UAWC \#2021-1120) and carried out in accordance with National Institutes of Health standards.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies
All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No | Yes
☒ ☐ Public health
☒ ☐ National security
☒ ☐ Crops and/or livestock
☒ ☐ Ecosystems
☒ ☐ Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No | Yes
☒ ☐ Demonstrate how to render a vaccine ineffective
☒ ☐ Confer resistance to therapeutically useful antibiotics or antiviral agents
☒ ☐ Enhance the virulence of a pathogen or render a nonpathogen virulent
☒ ☐ Increase transmissibility of a pathogen
☒ ☐ Alter the host range of a pathogen
☒ ☐ Enable evasion of diagnostic/detection modalities
☒ ☐ Enable the weaponization of a biological agent or toxin
☒ ☐ Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session
(e.g. UCSC)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument

Identify the instrument used for data collection, specifying make and model number.

Software

Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

| Cell population abundance | *Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.* |
| Gating strategy | *Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.* |

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI   ☐ Used   ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

| Model type and settings | *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).* |
| Effect(s) tested | *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.* |

Specify type of analysis: ☐ Whole brain   ☐ ROI-based   ☐ Both

| Statistic type for inference
(See Eklund et al. 2016) | *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.* |
| Correction | *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).* |

Models & analysis

n/a	Involved in the study
☐ ☐	Functional and/or effective connectivity
☐ ☐	Graph analysis
☐ ☐	Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.