Article

Striatum-wide dopamine encodes trajectory errors separated from value

https://doi.org/10.1038/s41586-025-10083-1

Received: 16 July 2024

Accepted: 19 December 2025

Published online: 11 February 2026

Check for updates

Eleanor H. Brown[1,2,3], Yihan Zi[1,2,3,4], Mai-Anh Vu[1,2], Safa Bouabid[1,2], Jack Lindsey[5,6,7], Chinyere Godfrey-Nwachukwu[1], Aaquib Attarwala[1,2], Ashok Litwin-Kumar[5,6], Brian DePasquale[2,4] & Mark W. Howe[1,2]✉

Goal-directed navigation requires animals to continuously evaluate their current direction and speed of travel relative to landmarks to discern whether they are approaching or deviating from their goal. Striatal dopamine release signals the reward-predictive value of cues[1,2], probably contributing to motivation[3,4], but it is unclear how dopamine incorporates an animal's ongoing trajectory for effective behavioural guidance. Here we demonstrate that cue-evoked striatal dopamine release in mice encodes bidirectional trajectory errors reflecting the relationship between the speed and direction of ongoing movement relative to optimal goal trajectories. Trajectory error signals could be computed from locomotion or visual flow, and were independent from simultaneous dopamine increases reflecting learned cue value. Joint trajectory error and cue-value encoding were reproduced by the reward prediction error term in a standard reinforcement learning algorithm with mixed sensorimotor inputs. However, these two signals had distinct state space requirements, suggesting that they could arise from a common reinforcement learning algorithm with distinct neural inputs. Striatum-wide multifibre array measurements resolved overlapping, yet temporally and anatomically separable, representations of trajectory error and cue value, indicating how functionally distinct dopamine signals for motivation and guidance are multiplexed across striatal regions to facilitate goal-directed behaviour.

Sensory cues experienced during goal-directed navigation serve as landmarks to direct appropriate adjustments in locomotor direction and speed (trajectory). Studies have implicated the striatum in landmark-based navigation strategies, which complement and compete with hippocampus- and cortex-dependent cognitive map-based strategies[5]. Accurately adjusting behaviour to an instructional cue requires information about whether the current trajectory is bringing an animal closer or further from the goal. For example, seeing a sign for an ice cream store while driving in one direction may indicate successful approach, whereas seeing it driving in the opposite direction would indicate an incorrect trajectory. In this case, the correct response (continue or turn back) depends on the relative motion of the landmark and observer, calculated from self-motion or visual flow. While representations underlying cognitive map-based navigation have been widely studied, the striatal signals guiding landmark-based trajectory adjustments remain unresolved.

Cues can trigger transient dopamine (DA) release in the striatum encoding a sudden change in the prediction of future reward, described in reinforcement learning models as reward prediction errors (RPEs)[1,6,7]. Cue-evoked RPEs scale with the expected reward probability, magnitude and timing, measures of the cue's motivational value[1,8,9]. Through bidirectional effects on striatal excitability and long-term plasticity[10], DA RPEs are positioned to influence motivation, associative learning and the vigour of ongoing or future behaviour[1,3,4,11]. However, in landmark-based navigation, a cue-evoked RPE signalling nearby reward would be insufficient for guiding trajectory adjustments without incorporating movement direction and speed relative to the goal.

Although RPE models can be adapted across contexts (or state spaces)[12,13], whether signals needed for trajectory adjustments are represented in striatal DA release is unclear. Most DA recordings come from instrumental tasks in which the animals are stationary at cue presentation, or Pavlovian tasks in which the cue and animal motion are irrelevant. In navigation tasks, ramping DA and place-value encoding reflect goal proximity[14–18] but, in these tasks, the animals always approached rewards along the correct trajectories, preventing the separation of trajectory accuracy from reward proximity.

Furthermore, when cues are encountered, the brain must disentangle potentially conflicting information about motivational value and behavioural accuracy. For example, seeing an ice cream store sign travelling in the wrong direction should elicit both a negative signal indicating movement away from the goal and a positive (value) signal indicating that reward is nearby. The negative signal could instruct corrective adjustments, whereas the positive component could regulate motivational drive. Recent theories propose that diverse RPEs are

[1]Department of Psychological & Brain Sciences, Boston University, Boston, MA, USA. [2]Centers for Systems Neuroscience and Neurophotonics, Boston University, Boston, MA, USA. [3]Graduate Program in Neuroscience, Boston University, Boston, MA, USA. [4]Department of Biomedical Engineering, Boston University, Boston, MA, USA. [5]Zuckerman Institute, Columbia University, New York, NY, USA. [6]Kavli Institute for Brain Science, Columbia University, New York, NY, USA. [7]Anthropic, San Francisco, CA, USA. ✉e-mail: mwhowe@bu.edu

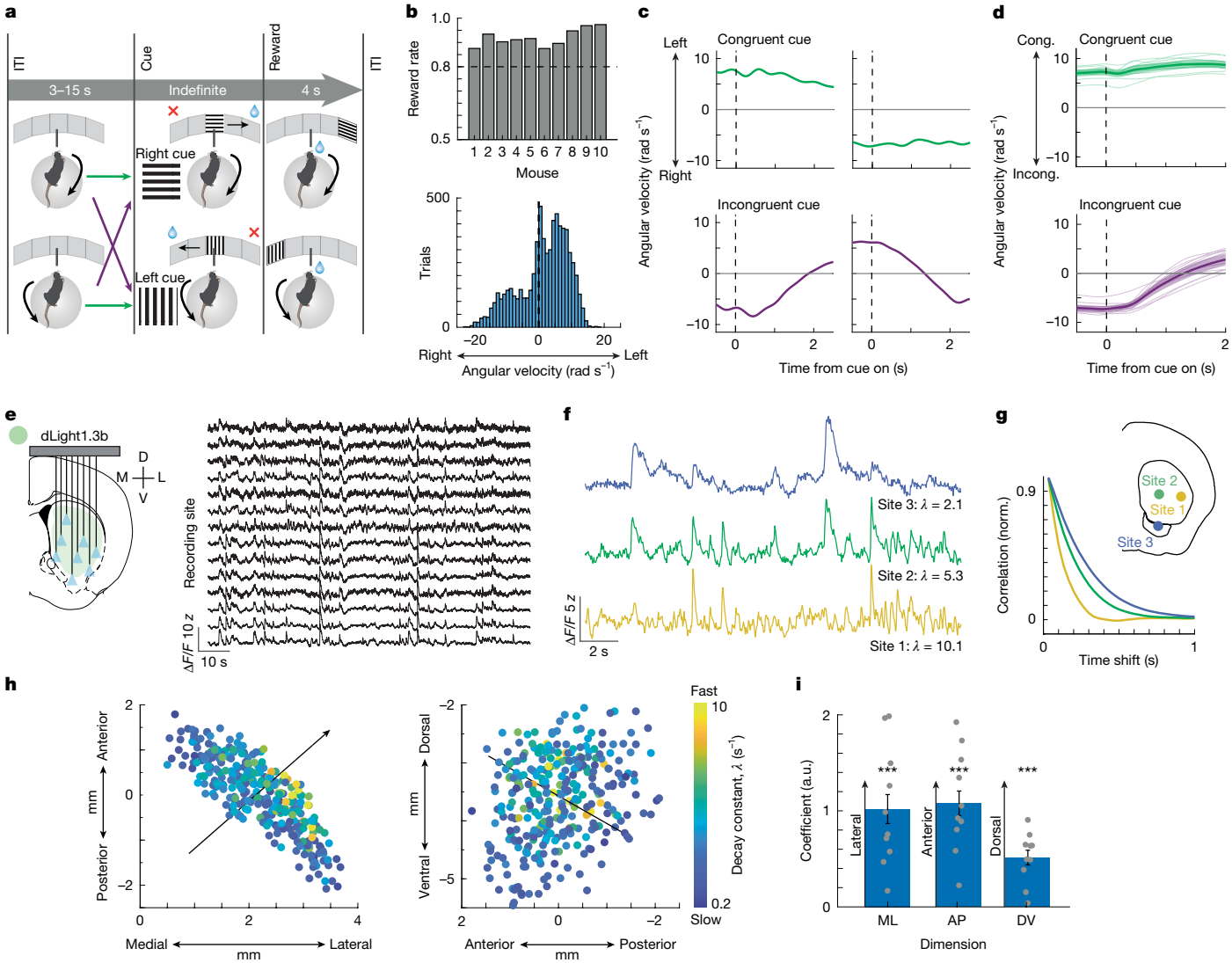


Fig. 1 | Striatum-wide DA release measurements during a visually guided instrumental task. a, Instrumental task schematic. Mice ran on a spherical treadmill during the ITI with angular velocities that were either congruent (green) or incongruent (purple) with the rewarded direction, indicated by a visual cue presented on each trial. The mouse image was adapted from ref. 61 under a CC BY 4.0 licence. **b**, The average reward rate across post-learning sessions for individual mice (top). Bottom, the average angular velocities before cue onset for each trial. **c**, The signed angular velocity on four example trials aligned to cue onset, with mice running left or right in congruent (top, green) or incongruent (bottom, purple) directions. **d**, The mean congruence-signed angular velocity aligned to the cue onset for congruent (cong.; top) and incongruent (incong.; bottom) trials. The thin lines represent individual mice. $n = 10$. **e**, Schematic of striatum-wide DA release measurements (left). Right, example z-scored $\Delta F/F$ traces recorded simultaneously in a single animal. D, dorsal; L, lateral; M, medial; V, ventral. **f**, Example traces from a single mouse, highlighting different temporal decay constants (λ). **g**, Autocorrelograms calculated across the entire recording period normalized to the maximum correlations for the three sites in **f**. Recording locations are shown in the coronal sections. **h**, Horizontal (top) and sagittal (bottom) maps of autocorrelation decay time constants for each site (circle) across mice. $n = 351$ total fibres, 10 mice. **i**, Spatial coefficients representing the magnitude and direction of the linear variation in the autocorrelation decay time constant along the AP, ML and DV axes. The grey dots show the coefficients computed for individual mice. Error bar centres represent the linear mixed-effects (LME) model coefficient estimates. ML: $P = 2.00 \times 10^{-6}$, $t_{312} = 5.53$; AP: $P = 1.14 \times 10^{-10}$, $t_{312} = 6.86$; DV: $P = 5.84 \times 10^{-10}$, $t_{312} = 6.58$; two-sided t-test on LME coefficients, post hoc Bonferroni corrected for the number of dimensions. $n = 316$ recording sites, 10 mice. The shaded regions in **d** and error bars in **i** show the 95% confidence intervals around the mean (**d**, centre line) and coefficient estimate (**i**, bar height). a.u., arbitrary units. For **e** and **g**, brain schematics were adapted from the Allen Mouse Brain Common Coordinate Framework (CCFv3) (https://atlas.brain-map.org/).

multiplexed across heterogeneous midbrain DA populations[12,19–22], but how distinct striatal regions read out these signals remains unresolved.

DA release in a visual instrumental task

We designed a task capturing key features of landmark-based navigation, in which mice experienced instructional cues during ongoing behaviour that indicated the running direction required for reward (Fig. 1a). Head-fixed mice ran on a floating ball, enabling two-dimensional (2D) locomotion[23]. On each trial, the mice were presented with one of two visual cues at the centre of a screen array, and the mouse's angular running velocity moved the cue across the screens. Trials ended when the cue moved laterally by around 33°. The median trial duration was 5 s from cue onset to reward, with an average of 111 trials per 33 min session. Each cue was presented randomly on 50% of trials and the cue pattern indicated the rewarded direction. Mice ($n = 10$) were trained to a reward rate of over 80% for at least three consecutive days (Fig. 1b). Importantly, before cue onset, the mice ran

with varying angular velocities (Fig. 1b). The cue identity indicated that travel was toward (congruent) or away (incongruent) from the rewarded direction. On correct congruent trials, the mice continued in the same direction, whereas, on correct incongruent trials, they switched (Fig. 1c,d). Thus, mice adjusted their ongoing locomotion rather than initiating actions from rest, analogous to encountering unexpected visual landmarks during navigation.

Encoding of RPE and behavioural variables in DA release is known to vary across striatal regions[24], so we applied a multifibre array approach[25] to measure DA release at many sites simultaneously (21–41 sites per mouse, 316 total across 10 mice) across the 3D striatum volume during task performance (Fig. 1e and Extended Data Fig. 1). This provided high signal-to-noise measurements of rapid DA release using the fluorescent sensor dLight1.3b[26] (Extended Data Fig. 2), isolated from artifacts using near-isosbestic illumination. Across task time, we found spatial gradients in intrinsic DA signal timescales, with the fastest decays in the anterior dorsolateral striatum, and slower decays in the ventromedial and posterior striatum (Fig. 1f–i). These results extend previous observations of dorsal–ventral (DV) differences in DA release kinetics and reuptake during behaviour[27,28].

DA release encodes trajectory errors

DA release increased at cue onset on both congruent and incongruent trials but was significantly higher on congruent trials at individual sites (Fig. 2a,b; 258 out of 316 sites total, mean 85.3 ± 12.4% per mouse congruent > incongruent; 2 out of 316 sites total, mean 0.7 ± 1.36% per mouse incongruent > congruent; $P < 0.05$, one-tailed Welch's *t*-test, Bonferroni corrected, $n = 10$ mice, mean ± 1.96 × s.e.m.), and across the population (Fig. 2d,e). With near-isosbestic 405 nm illumination, few sites showed significant trial-type differences (Fig. 2a,b; 15 out of 231 sites total, mean 11.1 ± 15.25% per mouse congruent > incongruent; 19 out of 231 sites, mean 12.3 ± 10.74% per mouse incongruent > congruent; $P < 0.05$, one-tailed Welch's *t*-test, Bonferroni corrected; $n = 7$ mice, mean ± 1.96 × s.e.m.), and no population-level effect was observed (Fig. 2d,e). In contrast to previously reported DA RPEs[8,9,29], these signals did not reflect fixed cue or action values but whether the animal's movement direction relative to the cue was bringing it closer to or further from the goal.

Mice ran with a range of speeds before the cue onset (Figs. 1b and 2f), so we tested whether congruence encoding scaled with speed. DA release at many sites scaled continuously with angular speed in opposite directions on congruent versus incongruent trials, with the largest signals on the fastest congruent trials and the smallest on the fastest incongruent (Fig. 2g). This scaling represents how correct or incorrect an animals' locomotion vector was at cue onset, which we term trajectory error (TE). To isolate TE effects, we modelled release at each timepoint after cue onset with finite impulse response (FIR) functions tiling a 2 s window, scaled trial-by-trial by task variables (for example, TE; Fig. 2h). Significant TE encoding was present at most sites but absent with near-isosbestic illumination (Fig. 2i,j; DA: 258 out of 316 sites total, 81.08 ± 11.13% per mouse, $n = 10$ mice, mean ± 1.96 × s.e.m.; isosbestic: 0 out of 231 sites, $n = 7$ mice; *t*-test on coefficients, $P < 0.05$ for ≥3 consecutive timepoints, Bonferroni corrected). Removing other behavioural variables reduced the model fit but did not alter TE encoding (Extended Data Fig. 3a–c). Replacing TE with a binary congruence term significantly reduced the model fit (Extended Data Fig. 3d) and yielded lower *t*-statistics for congruence coefficients (Extended Data Fig. 3e,f), confirming continuous TE scaling with speed. TEs did not differ across trial types split by congruence, running direction or cue identity (Extended Data Fig. 3g,h). Thus, cue-evoked DA release reflects a continuous measure of movement direction and speed relative to the optimal goal trajectory.

To test whether locomotion per se contributed to TE signalling, we first examined task-independent locomotion during the intertrial interval (ITI). DA release correlated positively with angular acceleration during contralateral running and negatively during ipsilateral running at 74 out of 193 sites (Extended Data Fig. 4a; 44.9 ± 3.7% sites significant, mean ± 1.96 × s.e.m.; *t*-test on model coefficients, Bonferroni corrected, $P < 0.05$, $n = 6$ mice). If acceleration alone explained TE scaling, TE encoding would be oppositely signed on contralateral and ipsilateral cue trials (Extended Data Fig. 4b). Instead, encoding was positive for both, with no significant magnitude differences (Extended Data Fig. 3g,h). TE coefficient magnitudes also did not differ between models with or without locomotor variables (Extended Data Fig. 3a–c). Finally, no velocity encoding occurred at cue onset in a Pavlovian task with stationary cues (0 out of 150 sites significant; $n = 4$ mice; Extended Data Fig. 4c–e). Thus, TE encoding was specific to instructional cues and their learned relationship with ongoing directional locomotion. In summary, although locomotion kinematics influence DA signalling as previously reported[19,30–32], they do not account for TE encoding.

TEs based on locomotor or visual input

During the task, angular velocity determined cue movement, so DA TE signals could arise from locomotor feedback, visual flow or both. To isolate visual flow, we trained naive mice ($n = 4$) on a Pavlovian task (visual-only (vis-only)) in which stimuli and reward contingencies matched the instrumental (locomotion + visual (loc+vis)) task, but cue speed and direction were yoked to loc+vis trials from other mice and independent of the subject's locomotion (Extended Data Fig. 5a). Mice rapidly learned the cue–direction–reward associations, pre-licking more on rewarded trials (Extended Data Fig. 5b,c). Visual TE (vis-TE), computed from cue motion, was uncorrelated with locomotor TE (loc-TE), based on the angular velocity at cue onset (Extended Data Fig. 5d,e; average Pearson's $r = 0.0089 \pm 0.026$, $P = 1.0$ for each mouse). Significant vis-TE, but not task-irrelevant loc-TE, encoding was present at most sites, with no magnitude difference from loc+vis TE (Extended Data Fig. 5g,h; 108 out of 116 sites, 92.71 ± 10.41% significant for vis-TE; 21 out of 116 sites, 16.79 ± 10.53% for loc-TE; *t*-test on model coefficients, $P < 0.05$ for ≥3 consecutive timepoints, Bonferroni corrected). Thus, TE computations can rely on visual flow alone, independent of locomotion.

We next tested whether DA TE encoding could be computed purely from locomotion, independent of visual flow. After loc+vis training, a subset of mice ($n = 2$) learned a locomotion-only (loc-only) version of the task in which cue–direction contingencies were identical, but the cue remained stationary after trial onset (Extended Data Fig. 5i). As in the loc+vis task, mice ran with a range of angular velocities before cue onset and adjusted locomotion according to cue–direction congruence, with no accuracy difference (Extended Data Fig. 5j–l). Significant TE encoding was observed at most sites (Extended Data Fig. 5m,n; 79 out of 79 and 83 out of 85 fibres significant; *t*-test on model coefficients, $P < 0.05$ for ≥3 consecutive timepoints, Bonferroni corrected). In summary, TE can be computed from locomotor feedback or visual flow independently, depending on task demands.

DA release encodes cue value

We next examined whether cue-evoked DA release encodes TE-independent cue value, consistent with canonical RPEs. We tested four RPE features. First, DA release should increase to the cue regardless of running direction. Our regression model captured a TE-independent cue-evoked DA increase at nearly all sites (non-TE; Extended Data Fig. 6a,c; 315 out of 316 sites, 99.66 ± 0.66% significant; $n = 10$ mice). Second, cue-evoked release should increase across early learning. In mice recorded across learning, cue-evoked DA magnitudes increased over early sessions (Extended Data Fig. 6b). Third, cue-value RPEs should update with recent reward history. Consistent with this, cue-evoked DA release was higher after rewarded than unrewarded trials (Extended Data Fig. 6e,f), and reward-history effects were captured

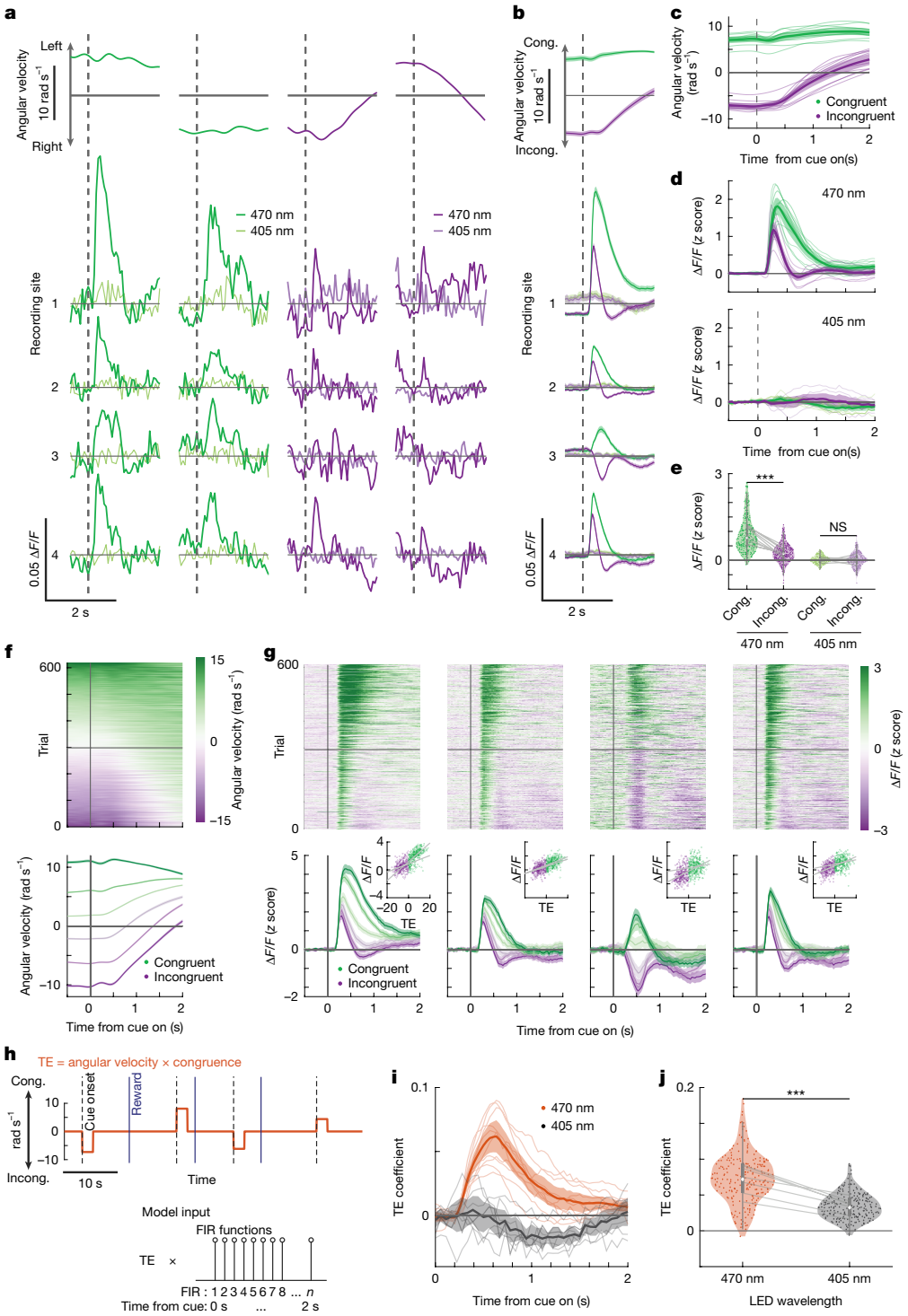


Fig. 2 | Cue-evoked DA release encodes trajectory errors. a, Signed angular velocity (+, left; −, right) (top) and Δ*F*/*F* (bottom) for four example fibres aligned to the cue onset on two congruent (green) and incongruent (purple) trials. Dark, 470 nm illumination; light, 405 nm illumination. **b**, Congruence-signed angular velocity (+, congruent; −, incongruent) averaged across rewarded trials for one mouse (top). Bottom, the corresponding Δ*F*/*F* for the four sites in **a**. $n = 321$ (congruent, 470 nm), $n = 297$ (incongruent, 470 nm), $n = 44$ (congruent, 405 nm), $n = 39$ (incongruent, 405 nm). **c**, Congruence-signed angular velocity, as in **b** (top), averaged across mice. $n = 10$. **d**, Δ*F*/*F*, as in **b** (bottom), averaged across sites and mice (LME model). Top, 470 nm, d.f. = 315, $n = 316$ sites; bottom, 405 nm, d.f. = 230, $n = 231$ sites, 7 mice. **e**, The mean Δ*F*/*F* within 1 s of cue onset for congruent and incongruent trials. The points show sites; the lines show mouse averages. Statistical analysis was performed using two-sided *t*-tests on LME model coefficients: $P = 2.0 \times 10^{-17}$, $t_{630} = -8.75$ (470 nm); $P = 0.96$, $t_{460} = 0.05$ (405 nm). **f**, The congruence-signed angular velocity ordered by TE (top) and TE tertile averages (bottom) for one mouse. **g**, Cue-aligned Δ*F*/*F* averages for four sites in the same animal, coloured by TE tertile. Inset: TE versus cue-evoked Δ*F*/*F* (1 s window). **h**, Schematic of the TE variable and FIR basis functions used in the regression model. **i**, Cue-aligned TE model coefficients averaged across sites and mice for 470 nm (orange) and 405 nm (grey). **j**, Mean TE model coefficients (across 1 s cue window) at each site for 470 nm ($n = 316$ sites, 10 mice) and 405 nm ($n = 231$ sites, 7 mice). $P = 4.26 \times 10^{-17}$, $t_{460} = -8.74$, $n = 231$ sites, 7 mice, two-sided *t*-test on LED wavelength effect. The shaded regions in **b**–**d**,**f**,**g**,**i** show the 95% confidence intervals around the mean (centre line). The thin lines in **c**–**e**,**i**,**j** show the mouse averages. For the box plots, the centre line shows the median, the box bounds show the first and third quartiles, and the whiskers show extremes values <1.5× interquartile range from the bounds.

in our regression model (Extended Data Fig. 6d; 107 out of 261 fibres, 42.98 ± 17.23% significant; $n = 8$ mice, mean ± 1.96 × s.e.m.). Fourth, cue-value RPEs should discriminate cues with different relative reward rates and latencies. Mice showed inherent (typically contralateral) running biases (Methods and Extended Data Fig. 6j), resulting in shorter latencies to reward and slightly higher reward rates for bias cues (Extended Data Fig. 6k,l). Cue-evoked DA release was higher for these higher-valued contralateral bias cues (Extended Data Fig. 6h,i). This effect was captured in the regression model as positive cue-identity coefficients (223 out of 316 fibres, 72.79 ± 17.21% significant; $n = 10$ mice, mean ± 1.96 × s.e.m.) and coefficient magnitudes correlated with a contralateral bias index (Extended Data Fig. 6g). Thus, in addition to TEs, cue-evoked DA release encodes learned, temporally discounted cue–reward associations, consistent with Pavlovian RPEs and captured independently from TE by the non-TE, reward-history and cue-identity coefficients.

Modelled TEs and cue-value RPEs

Reinforcement learning models posit that DA release to cues encodes RPEs emerging from updates to a value function representing temporally discounted expected future rewards[1,6,7,33]. In our task, TE at cue onset was inversely related to log-transformed time to reward (Fig. 3a; Pearson's $r = -0.41 \pm 0.12$, $n = 10$ mice, mean ± 1.96 × s.e.m.; $P < 0.05$ for 9 out of 10 mice, Bonferroni corrected), suggesting that DA TE encoding could be modelled within a temporal-difference reinforcement learning (tdRL) framework using an actor–critic architecture[6,33] (Fig. 3). As TEs incorporate cue identity and continuous movement, we used a state space with a single-layer feedforward network in which units encoded mixed representations of randomly weighted sensorimotor variables[34] (Fig. 3b and Methods). This representation provided input to the critic (for RPE computations) and the actor (for action selection). The model reproduced accurate choice behaviour and replicated DA TE encoding (Fig. 3c–f), indicating that TEs can be modelled as RPEs computed from mixed sensorimotor inputs.

We next examined whether the RPE term encoded a learned, TE-independent cue value. The RPE term increased across early trials and was sensitive to the outcome of the previous trial (Fig. 3g–i), consistent with our DA data (Extended Data Fig. 6). Incorporating a directional running bias based on mouse behaviour yielded a larger RPE for the bias-associated cue, due to faster rewards and a higher rate for that cue (Fig. 3j and Extended Data Fig. 7a–f). Thus, with a mixed-sensorimotor state space, the RPE term simultaneously encodes TE and a learned, movement-independent cue value.

We next evaluated modified state spaces to test the importance of specific variables and mixed selectivity for generating TEs (Fig. 3k–n and Extended Data Fig. 7g–l). Simulations without cue identity performed at chance and were excluded. Removing position or velocity inputs reduced TE encoding and removing both abolished it (Fig. 3k and Extended Data Fig. 7g). Eliminating the feedforward layer, producing units selective for single variables, further reduced TE encoding relative to all mixed-input models (Fig. 3k and Extended Data Fig. 7j). By contrast, measures of cue value were preserved across all mixed and unmixed state spaces containing cue identity (Fig. 3l–n and Extended Data Fig. 7g–l). Thus, while cue inputs alone are sufficient for RPE encoding of learned cue–reward associations, TE encoding requires a state space with mixed representations of relevant sensorimotor variables.

TEs and cue value RPEs are separable

Our results indicate that cue-evoked DA release simultaneously represents RPEs encoding learned cue value and a movement-dependent TE, with distinct state–space requirements. We examined how these computationally and functionally distinct signals are separated in the striatum. To test temporal separability, we compared the time courses of DA TE and cue-value (non-TE, reward history, cue identity) encoding (Fig. 4a–c). Cue-value variables were highly overlapping, with no significant lags in maximal cross-correlations, indicating a common cue-value RPE computation (Fig. 4c). By contrast, significant lags existed between TE and each cue-value variable, with cue value preceding TE (Fig. 4a,b). Thus, cue-value and TE signals are temporally separable, consistent with an RPE model in which cue inputs are processed faster than mixed sensorimotor inputs[2].

We next used our high-density multifibre array recordings to examine striatum-wide spatial patterns of TE and cue value encoding. TE encoding maps revealed a 3D gradient that extended across the striatum volume (Fig. 4d). A linear regression on TE coefficients by spatial position revealed a significant anterior–posterior (AP) gradient, with the strongest TE encoding in the anterior striatum (Fig. 4e). By contrast, all cue-value variables exhibited significant medial–lateral (ML) gradients but not AP gradients, with strongest encoding in the medial striatum (Fig. 4g,h). From these coefficients, we constructed 3D gradient vectors, capturing axes of maximal variation (Fig. 4f,i). Gradient vectors for TE and each cue-value variable were consistent and significantly concentrated across animals (Fig. 4f,i). Comparisons across cue-value variables revealed small but significant DV differences (Fig. 4h and Extended Data Fig. 8a), but gradient vector directions did not differ significantly (Fig. 4i), indicating a shared axis of variation for RPEs based on movement-independent cue value.

To test for anatomical separability of TE and cue value encoding, we quantitatively compared their spatial gradients. TE spatial coefficients differed significantly from each cue value variable in the ML and AP dimensions (Fig. 4j), and their 3D gradient vectors were significantly different (Fig. 4k). Notably, the TE and cue value gradient vectors were approximately orthogonal (Fig. 4l), enabling distributed and overlapping, yet separable, representations. Together, these results support a spatiotemporal model of cue-evoked DA release in which computationally distinct RPEs are separated in time and vary along orthogonal anatomical axes, providing a mechanism for downstream decoding.

Spatial multiplexing of DA signals

Cue-value and TE signals transmit distinct information about reward availability and movement toward or away from reward. Our results show that these signals are spatially and temporally separated, enabling location-specific cue-evoked DA release determined by the summation of cue-value and TE components (Fig. 5a,b). Net DA release at each site reflects the local weights of these components (Fig. 4d–l and Extended Data Fig. 8b), their temporal offset (around 0.3 s; Fig. 4a–c) and spatially varying signal-decay constants (Fig. 1h,i). To illustrate this, we focused on incongruent trials in which cue value and TE signals were diametrically opposed (positive cue value, negative TE). Slower decays increased overlap between the DA increase (cue value) and the DA dip (TE), resulting in reduced dip magnitudes. Faster decays produced less summation, greater component separability and larger dips (Fig. 5c–e). A comparison of sites with similar decays showed that relative cue-value and TE weights shape the DA peak and dip magnitudes (Fig. 5f–h). To capture the net signal summation across sites, we compared the spatial patterns of the maximum and minimum DA $\Delta F/F$ magnitudes. Cue-evoked peaks and dips were expressed across distinct, nearly orthogonal spatial gradients, with the largest dips in the anterior dorsolateral striatum and the largest peaks in the anterior dorsomedial striatum (aDMS; Fig. 5i–k and Supplementary Table 1). The gradients resembled, but were distinct from, those defined by the cue value and TE model coefficients (compare Fig. 5j,k with Fig. 4d–l). Notably, the dips showed a significant DV gradient, absent in TE magnitudes (compare Fig. 5j with Fig. 4e), probably reflecting faster decays in the anterior dorsolateral striatum relative to the ventral striatum (Fig. 1h,i). In summary, intrinsic kinetics and the spatiotemporal separation of cue

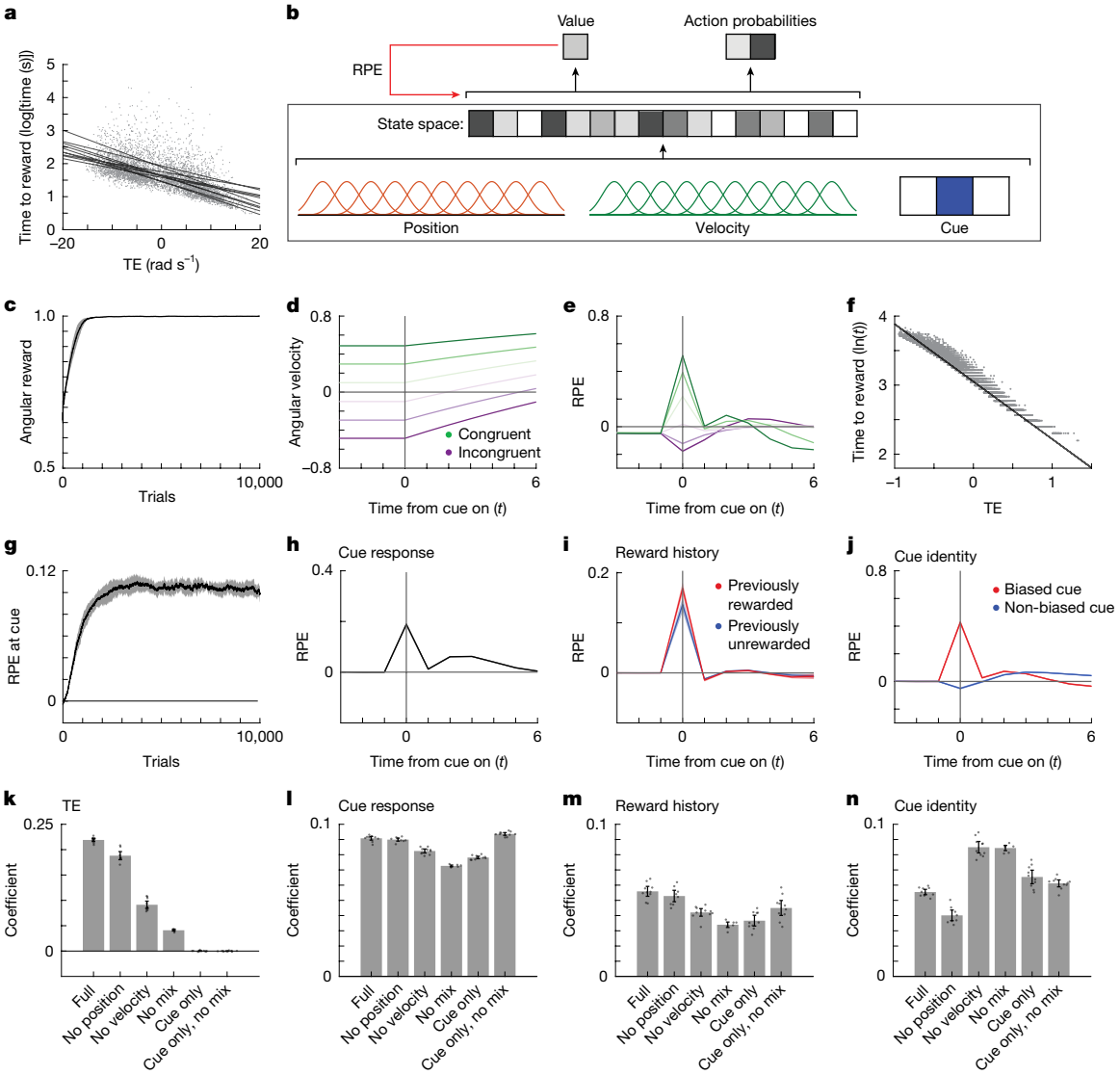


Fig. 3 | Distinct state space requirements for reproducing TE and cue-value encoding in a tdRL algorithm. a, TE versus log-transformed time to reward for all trials (grey dots) across mice. The grey lines represent least-squares regression lines for individual mice. **b**, Schematic of the temporal tdRL algorithm with a high-dimensional state space with random nonlinear mixing of relevant sensorimotor variables (Methods). **c**, tdRL performance (reward rate) across trials in a virtual task replicating our task in mice (Fig. 1a), smoothed with a 500-trial sliding window and averaged across $n = 10$ simulations. **d**, Simulated angular velocity aligned to the cue onset, signed by congruence, for congruent (green) and incongruent (purple) trials, split into thirds by velocity magnitude and averaged across 10 simulations. **e**, RPE aligned to cue onset, averaged as in **d**. **f**, The trial-by-trial relationship between TE and log-transformed time to reward for ten simulations. The grey lines are the least-squares regression lines. **g**, RPE at cue onset ($t = 0$), smoothed with a 500-trial sliding window, averaged across $n = 10$ simulations, as in **c**. **h**, The cue-onset-aligned RPE across trials after stable performance averaged across $n = 10$ simulations. **i**, The cue-onset aligned RPE as in **h** across trials split by reward history (red, following rewarded; blue, following unrewarded) averaged across $n = 10$ simulations. **j**, The cue-onset-aligned RPE as in **h** split by cue identity (red, biased direction cue; blue, non-biased direction cue) averaged across $n = 10$ simulations. **k**, Average RPE encoding of TE at the cue onset (averaged across $t = 0$,1 and baseline corrected) for each tested state space, determined from linear model coefficients across $n = 10$ simulations. **l**–**n**, Average encoding of non-TE cue increases (**l**), reward history (**m**) and cue identity (**n**) in the RPE model term ($n = 10$ simulations), as in **k**. The shaded regions in **c**–**e**,**g**–**j** and error bars in **k**–**n** show the 95% confidence intervals around the mean (center line, **c**–**e**,**g**–**j**) and coefficient estimates (bar height, **k**–**n**).

value and TE encoding enable multiplexing of functionally distinct DA signals across striatal regions.

To investigate the relationships between TE signals and online or future behaviour adjustments, we compared DA dips (negative TEs) on correct (switch) and incorrect (no switch) incongruent trials (Extended Data Fig. 9). Some sites showed larger dips on correct switches, others on incorrect trials (Extended Data Fig. 9a,b). These differences were anatomically organized, with larger dips on incorrect trials in the anterior ventral striatum and larger dips on correct switch trials in the dorsolateral striatum (DLS; Extended Data Fig. 9c–e). Thus, DLS signals better track ongoing locomotion adjustments, whereas ventral striatum dips more strongly signal incorrect trajectories, potentially supporting future behavioural updating.

To further constrain behavioural functions of TEs, we recorded DA release across learning in a subset of mice ($n = 3$). A dynamic regression model[35] estimated the TE evolution on single trials across learning (Extended Data Fig. 10). TE coefficients increased rapidly and reached statistical significance before performance exceeded chance levels, suggesting a role in learning (Extended Data Fig. 10a,b). Although early TE encoding emergence was widespread, the time course differed

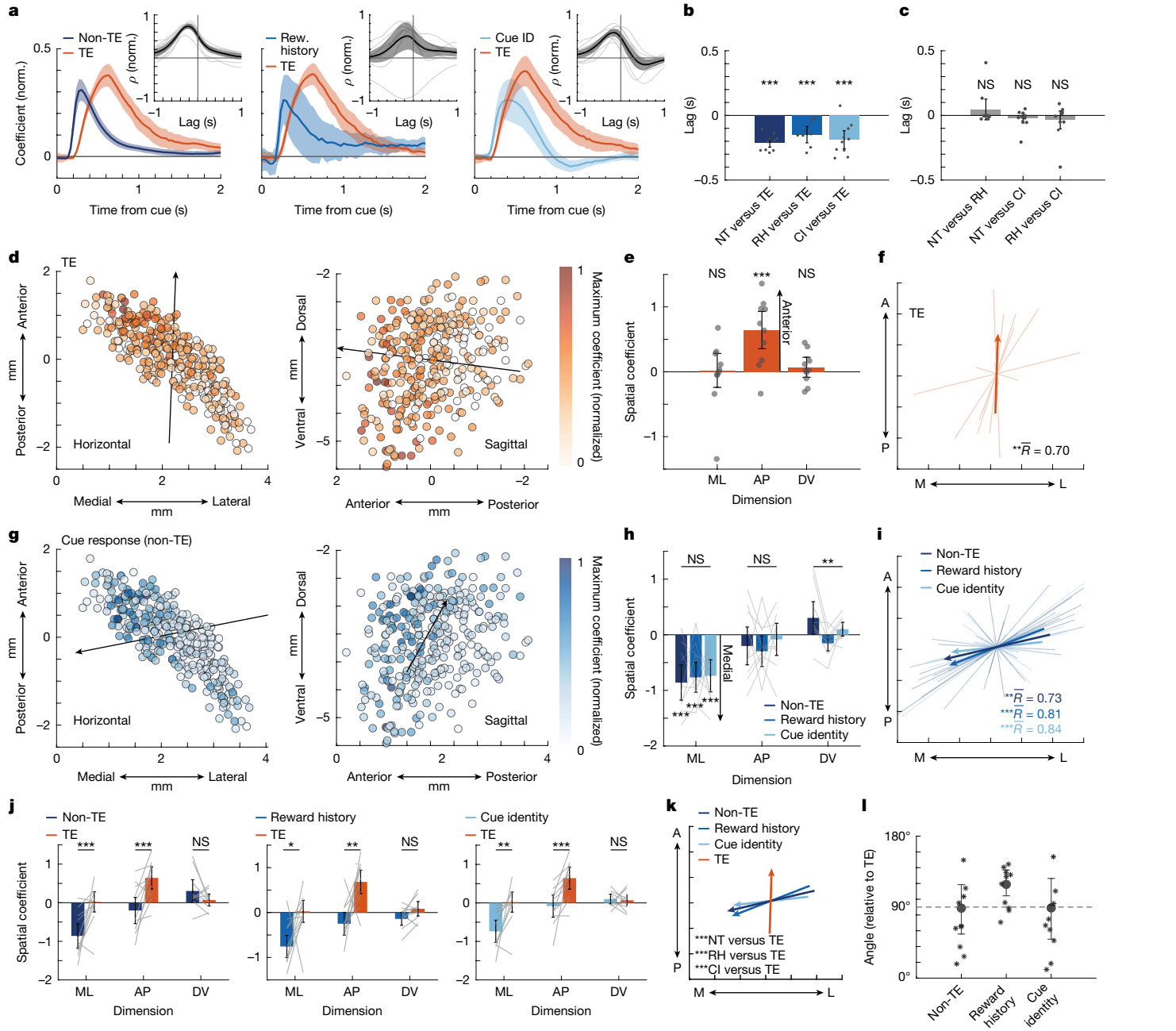


Fig. 4 | DA encoding of TE and cue value are separated in time and space. **a**, Cue-aligned model coefficients for cue value (blue) and TE (orange), averaged across sites and mice with significant encoding. Insets: cross-correlations (normalized (norm.) to autocorrelation at lag 0) between cue-value and TE coefficients. Negative lags, cue-value preceding TE. Rew., reward. **b**, Lags to peak correlations between TE and cue-value variables, averaged across sites and mice with significantly positive correlations (Supplementary Table 1). CI, cue identity; NT, non-TE; RH, reward history. **c**, Lags to peak correlations between each pair of cue-value coefficients, as in **b** (Supplementary Table 1). **d**, Horizontal (left) and sagittal (right) maps of maximum TE coefficient magnitudes for each site (n = 316 sites, 10 mice), normalized to the global maximum. The blank circles indicate non-significant coefficients. The arrows indicate the axis of maximal variation based on spatial coefficients in **e**. **e**, Spatial coefficients describing linear variation in TE magnitude (two-sided t-test on LME model coefficients ML: $t_{312} = 0.16$, $P = 1.0$; AP: $t_{312} = 4.40$, $P = 4.55 \times 10^{-5}$; DV: $t_{312} = 0.88$, $P = 1.0$; Bonferroni corrected). **f**, Total TE gradient vector based on the coefficients in **e**. R, the mean resultant length. TE: $S^* = 16.00$, $P = 0.0011$, weighted 3D Rayleigh test. **g**, Maps of the maximum non-TE coefficients (TE-independent cue responses), as in **d**. **h**, Spatial coefficients describing linear variation in each cue-value coefficient magnitude (ML: $F_{2,936} = 0.34$, $P = 1.0$; AP: $F_{2,936} = 0.74$, $P = 1.0$; DV: $F_{2,936} = 6.38$, $P = 0.0018$; Bonferroni-corrected ANOVA on location × variable type; individual effects, Supplementary Table 1). **i**, The total gradient vector for the cue-value coefficients in **h**; each vector significantly concentrated (Supplementary Table 1) with no difference in mean directions ($F_{4,54} = 1.84$, $P = 0.13$, $n = 10$). **j**, Comparison of TE (orange) and cue-value (blue) coefficient magnitudes (t-tests on location or location × coefficient type effects in the LME model; n = 316 sites; Supplementary Table 1). **k**, TE versus cue-value gradient vectors (two-sample high-concentration F-test). **l**, The angle between TE and cue-value gradient vectors averaged across mice. n = 10. The error bars in **b**,**c**,**e**,**h**,**j**,**l** and shaded regions in **a** show the 95% confidence intervals around the coefficient estimates (centre lines in **a** and bar heights in **b**,**c**,**e**,**h**,**j**). The thin lines in **a**,**f**,**h**–**j** and points in **b**,**c**,**e**,**l** represent individual mice. ***$P < 0.001$, **$P < 0.01$.

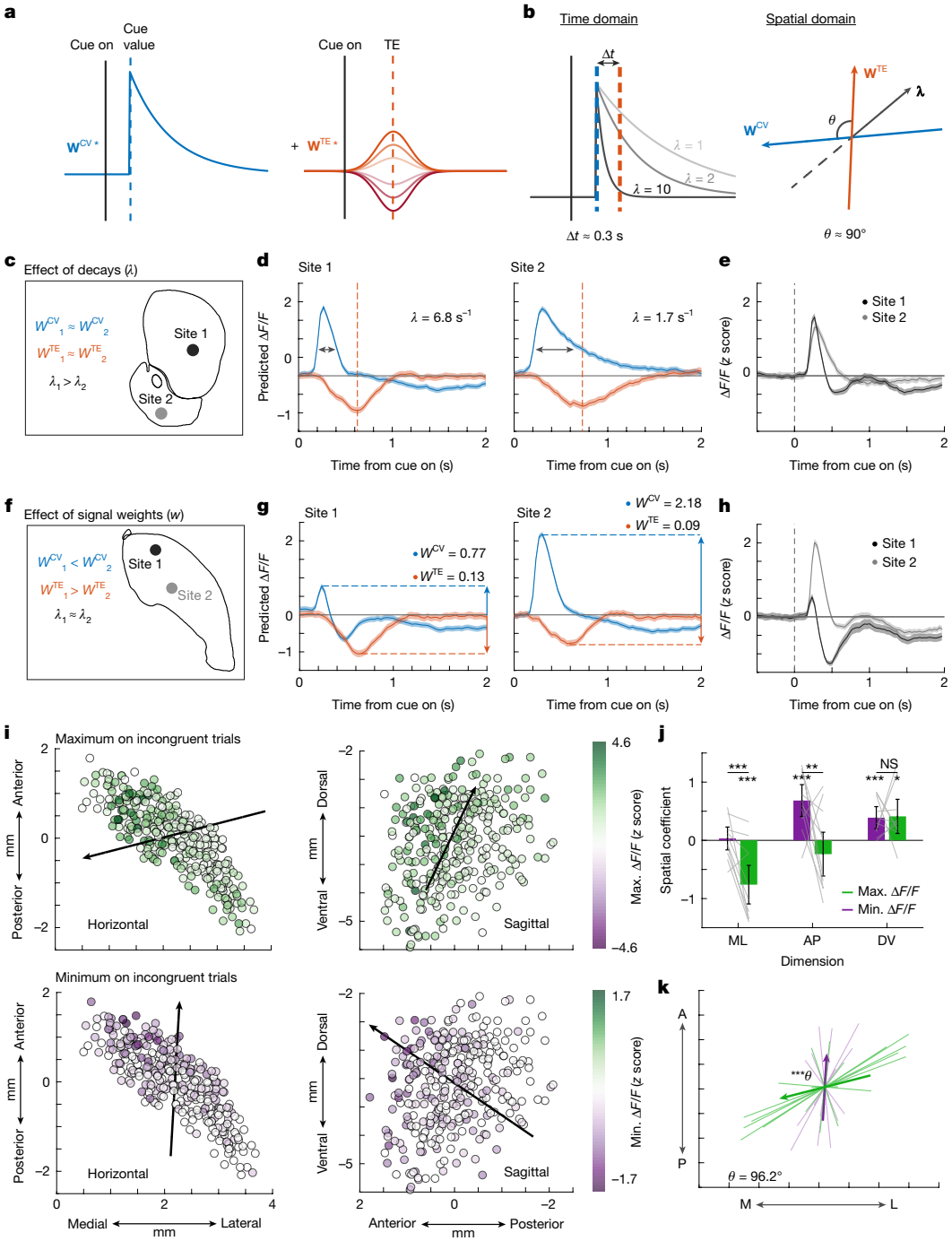


Fig. 5 | Orthogonal cue-value and TE gradients and intrinsic decay kinetics determine region-specific DA release. a, Schematic of cue-value and TE components with different location-specific weights ($\mathbf{W}^{CV}$ and $\mathbf{W}^{TE}$), estimated by model coefficients). **b**, Schematic of the separation of cue value and TE encoding in time (left) and space (right). Left, the dashed lines show the mean timing of maximal non-TE (blue) and TE (orange) coefficients. The grey lines show the effects of regional differences in intrinsic signal decay time constants (λ) on the temporal separation of cue-value and TE components. Right, spatial gradients for cue-value and TE ($\mathbf{W}^{CV}$ and $\mathbf{W}^{TE}$) and decay constants ($\boldsymbol{\lambda}$) jointly determine component summation across the striatum. **c**, Coronal locations for two example sites with a similar W^{CV} and W^{TE} but different λ. **d**, Predicted cue-aligned $\Delta F/F$ for non-TE (blue) and TE (orange) components on incongruent trials for sites in **c**. The dashed lines mark DA dip troughs. **e**, The observed cue-aligned $\Delta F/F$ (z scored) for the same sites, showing that a slower decay increases temporal overlap and reduces dip depth. **f**, The horizontal locations for two sites with similar λ but different W^{CV} and W^{TE}. **g**, The predicted $\Delta F/F$ for cue value and TE on incongruent trials for sites in **f**. **h**, The corresponding observed $\Delta F/F$. **i**, Horizontal (left) and sagittal (right) maps of maximum (top) and minimum (bottom) cue-evoked $\Delta F/F$ on incongruent trials ($n = 316$ sites, 10 mice). The empty bubbles indicate no significant difference versus bootstrapped null. **j**, Spatial coefficients quantifying variation in $\Delta F/F$ minima and maxima (ML: $t_{624} = 4.0$, $P = 2.23 \times 10^{-4}$; AP: $t_{624} = 3.14$, $P = 0.0053$; DV: $t_{624} = -0.16$, $P = 1.0$; two-sided t-test on the LME interaction term, Bonferroni corrected; individual effects, Supplementary Table 1). The thin lines represent coefficients for individual mice. **k**, Total gradient vectors for $\Delta F/F$ minima and maxima; both are significantly concentrated (Rayleigh test; Supplementary Table 1) with directions differing significantly ($F_{2,38} = 8.37$, $P = 0.001$, two-sample, one-sided, high-concentration F-test). The shaded regions in **d**,**e**,**g**,**h** and error bars in **j** show the 95% confidence intervals around the FIR coefficient estimates (**d** and **g**), mean $\Delta F/F$ (**e** and **h**) or LME coefficients (**j**). For **c** and **f**, brain schematics were adapted from the Allen Mouse Brain Common Coordinate Framework (CCFv3) (https://atlas.brain-map.org/).

across striatal regions. Non-negative matrix factorization was used to identify early and late components to the TE time course, revealing a spatial bias towards early increases in the anterior medial striatum and later increases in the lateral striatum (Extended Data Fig. 10e–h). This organization was consistent across mice (Extended Data Fig. 10f–h), and resembled dip gradients on correct and incorrect incongruent trials (compare Extended Data Fig. 10e–h with Extended Data Fig. 9c–e). Together, these results suggest that DA TEs may support early learning of task contingencies in medial striatum and later behavioural refinement in lateral striatum, consistent with known region-specific functions[36,37].

Discussion

Decades of work have established that the DA system can signal RPEs reflecting the motivational value of cues[1–3,9]. Here we expand this framework in two ways. First, we report that, in addition to cue-value RPEs, cue-evoked DA release encodes TEs that signal whether, and how fast, animals are travelling in the correct or incorrect direction relative to a goal (Fig. 2). TE encoding can be referenced to either locomotion, visual flow or both, allowing flexibility when only certain reference frames are relevant (Extended Data Fig. 5). This signal represents information critical for evaluating ongoing locomotor trajectories during landmark based navigation. Second, TEs are separable from cue-value RPEs in time and anatomical space, enabling decoding of multiplexed DA signals, even when value and TE signals conflict (Figs. 4 and 5). TEs can be reproduced by an RPE term in a standard tdRL model but require state–space representations distinct from cue-value based RPEs (Fig. 3), suggesting a shared algorithm operating on DA neuron populations receiving distinct combinations of sensorimotor inputs.

Error signals encoding mismatches between predicted sensory feedback and motor commands occur across multiple species and brain regions and support roles in online error correction and learning[38–40]. For example, in zebra finches, DA neurons encode performance errors in vocal motor output relative to a tutor template[39]. By contrast, TEs do not indicate motor errors relative to a fixed sensory or motor template, but reflect flexible associations between cues and ongoing motion relative to learned goals. This property is essential for linking landmarks to goal-directed trajectories through learning and to guide appropriate online course corrections. Signals encoding negative errors in locomotion direction have been reported in a sparse subpopulation of posterior parietal cortex interneurons during navigation; however, it is unclear how they could broadly impact learning or guide movement[41]. TE encoding in DA release was widespread across the anterior striatum (Fig. 4d) and is therefore positioned to powerfully impact striatal output and behaviour across multiple parallel basal ganglia circuits.

The tdRL framework has been applied to explain the role of hippocampal and striatal networks in learning associations between locations, actions and outcomes to guide navigation[5,42,43]. However, to our knowledge, DA TE encoding has not been experimentally observed or formally predicted from tdRL models. While the tdRL framework specifies a common learning algorithm, a nearly infinite set of possible state–space representations can generate many distinct RPE signals. Our simulations showed that both TE and cue-value encoding could be reproduced as RPEs generated through tdRL (Fig. 3). However, whereas cue-value RPEs could be generated with cue identity inputs alone, TE signals required mixed representations of cue identity, velocity and position (Fig. 3k–n). Such representations could arise from striatum and midbrain-projecting cortical neurons with mixed selectivity to locomotor and sensory variables[34,44]. Although our findings fit within a tdRL framework, other learning algorithms are possible[45]. Rather than endorsing any specific learning algorithm, our results underscore the utility of expanding conjunctive model state spaces to capture dynamics associated with naturalistic behaviours[13].

Individual midbrain DA neurons can represent different information decodable at the level of neuronal subpopulations[19–21,30,46], but it is unclear how diverse DA signals are decoded by downstream striatal regions, given the large, overlapping axonal arbours of DA neurons[47]. Several reports of regional heterogeneities in striatal DA signals have contrasted the expression of RPEs and non-RPEs[30,48–50], while others have reported heterogeneity in the presence and magnitude of RPE signals within a single dimension (for example, value) across select subregions[51–53]. Limitations in the spatial coverage, density and resolution of previous DA recording methods have fostered a compartmentalized perspective in which particular DA signal profiles are broadly associated with large striatum subregions. With our high-density arrays, we obtained a global view of striatum-wide DA dynamics, unconstrained by preselected targets, revealing functional signal gradients spanning large-scale 3D volumes. Gradients of cue-value and TE encoding overlap extensively, raising the question of how these signals are disambiguated downstream, particularly when they conflict. Despite joint encoding at individual locations, the signals were separable in time and anatomical space. Cue-evoked value encoding preceded TE encoding creating a temporal separation shaped by spatially varying intrinsic DA decay time constants (Fig. 5). Moreover, orthogonal spatial gradients in relative encoding magnitudes (Fig. 4k,l) enabled unique summation of TE and cue value signals at each location, defining regional DA release specialized for distinct aspects of landmark-based navigation (behavioural adjustments and general motivation/vigour, respectively; Fig. 5). These results suggest general spatial and temporal neural mechanisms for multiplexing RPE signals with distinct inputs and functions. Future studies will be needed to test whether these same principles apply to other dimensions of DA signalling.

Regarding how TE signals might influence behaviour and learning during goal-directed navigation, the role of striatal DA in controlling behaviour is debated, with some studies indicating a role in subsecond movement control and others an exclusive role in learning[30,32,54–57]. Artificially augmenting DA can modulate striatal neurons on rapid (hundreds of ms) and long term (days) timescales[10,58]. TEs could therefore serve dual roles in modulating both ongoing choice (continue or adjust locomotor trajectory) and associative cue learning. While direct tests of causal relationships will require precisely timed manipulations calibrated to endogenous distributions over large volumes[25], our results constrain possible region-specific functional roles. First, on incongruent trials, DA increases reflecting cue value were strongest in the aDMS, while dips encoding the negative TE were strongest in the anterior DLS (aDLS; Fig. 5i–k). Thus, aDMS DA is positioned to modulate the general motivation to work, whereas aDLS DA is positioned to modulate behavioural adjustments, roles compatible with previous regional findings[11,14,30,57]. Second, the emergence of TEs preceded successful task performance, indicating that TEs are not sufficient alone to drive movement changes, but may contribute to learning and/or combine with other circuit dynamics to modulate performance (Extended Data Fig. 10a,b). Third, the medial striatum had stronger TE encoding before learning with larger dips when mice continued an incorrect trajectory, while the lateral striatum had stronger TE encoding stabilized performance, with larger dips on correct switches (Fig. 5). Thus, lateral TEs are better positioned to influence the immediate switching probability, while medial TEs are better positioned for long-term updating. Cue-evoked DA dips in aDLS have been recently linked to acetylcholine increases in the same region during Pavlovian extinction learning[59], suggesting a potential role in updating behaviour to facilitate behavioural flexibility[60].

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/s41586-025-10083-1.

1. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
2. Schultz, W. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* **30**, 259–288 (2007).
3. Berridge, K. C. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology* **191**, 391–431 (2007).
4. Wise, R. A. Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**, 483–494 (2004).
5. Chersi, F. & Burgess, N. The cognitive architecture of spatial navigation: hippocampal and striatal contributions. *Neuron* **88**, 64–77 (2015).
6. Joel, D., Niv, Y. & Ruppin, E. Actor–critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw.* **15**, 535–547 (2002).
7. Montague, P., Dayan, P. & Sejnowski, T. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
8. Tobler, P. N., Fiorillo, C. D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
9. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
10. Gerfen, C. R. & Surmeier, D. J. Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* **34**, 441–466 (2011).
11. Saunders, B. T., Richard, J. M., Margolis, E. B. & Janak, P. H. Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nat. Neurosci.* **21**, 1072–1083 (2018).
12. Lee, R. S., Sagiv, Y., Engelhard, B., Witten, I. B. & Daw, N. D. A feature-specific prediction error model explains dopaminergic heterogeneity. *Nat. Neurosci.* https://doi.org/10.1038/s41593-024-01689-1 (2024).
13. Gershman, S. J. & Daw, N. D. Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annu. Rev. Psychol.* **68**, 101–128 (2017).
14. Hamid, A. A. et al. Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**, 117–126 (2016).
15. Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M. & Graybiel, A. M. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* **500**, 575–579 (2013).
16. Kim, H. R. et al. A unified framework for dopamine signals across timescales. *Cell* **183**, 1600–1616 (2020).
17. Krausz, T. A. et al. Dual credit assignment processes underlie dopamine signals in a complex spatial environment. *Neuron* **111**, 3465–3478 (2023).
18. Farrell, K., Lak, A. & Saleem, A. B. Midbrain dopamine neurons signal phasic and ramping reward prediction error during goal-directed navigation. *Cell Rep.* **41**, 111470 (2022).
19. Engelhard, B. et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* **570**, 509–513 (2019).
20. Dabney, W. et al. A distributional code for value in dopamine-based reinforcement learning. *Nature* **577**, 671–675 (2020).
21. Sousa, M. et al. A multidimensional distributional map of future reward in dopamine neurons. *Nature* **642**, 691–699 (2025).
22. Masset, P. et al. Multi-timescale reinforcement learning in the brain. *Nature* **642**, 682–690 (2025).
23. Dombeck, D. A., Harvey, C. D., Tian, L., Looger, L. L. & Tank, D. W. Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat. Neurosci.* **13**, 1433–1440 (2010).
24. Collins, A. L. & Saunders, B. T. Heterogeneity in striatal dopamine circuits: form and function in dynamic reward seeking. *J. Neurosci. Res.* **98**, 1046–1069 (2020).
25. Vu, M.-A. T. et al. Targeted micro-fiber arrays for measuring and manipulating localized multi-scale neural dynamics over large, deep brain volumes during behavior. *Neuron* **112**, 909–923 (2024).
26. Patriarchi, T. et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* **360**, eaat4422 (2018).
27. Mohebi, A., Wei, W., Pelattini, L., Kim, K. & Berke, J. D. Dopamine transients follow a striatal gradient of reward time horizons. *Nat. Neurosci.* **27**, 737–746 (2024).
28. Jørgensen, S. H. et al. Behavioral encoding across timescales by region-specific dopamine dynamics. *Proc. Natl Acad. Sci. USA* **120**, e2215230120 (2023).
29. Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* **9**, 1057–1063 (2006).
30. Howe, M. W. & Dombeck, D. A. Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature* **535**, 505–510 (2016).
31. Dodson, P. D. et al. Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proc. Natl Acad. Sci. USA* **113**, E2180–E2188 (2016).
32. Da Silva, J. A., Tecuapetla, F., Paixão, V. & Costa, R. M. Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* **554**, 244–248 (2018).
33. Sutton, R. S. Learning to predict by the methods of temporal differences. *Mach. Learn.* **3**, 9–44 (1988).
34. Lindsay, G. W., Rigotti, M., Warden, M. R., Miller, E. K. & Fusi, S. Hebbian learning in a random network captures selectivity properties of the prefrontal cortex. *J. Neurosci.* **37**, 11021–11036 (2017).
35. Roy, N. A., Bak, J. H., Akrami, A., Brody, C. D. & Pillow, J. W. Extracting the dynamics of behavior in sensory decision-making experiments. *Neuron* **109**, 597–610 (2021).
36. Graybiel, A. M. Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* **31**, 359–387 (2008).
37. Yin, H. H. & Knowlton, B. J. The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* **7**, 464–476 (2006).
38. Shadmehr, R., Smith, M. A. & Krakauer, J. W. Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* **33**, 89–108 (2010).
39. Gadagkar, V. et al. Dopamine neurons encode performance error in singing birds. *Science* **354**, 1278–1282 (2016).
40. Greenstreet, F. et al. Dopaminergic action prediction errors serve as a value-free teaching signal. *Nature* https://doi.org/10.1038/s41586-025-09008-9 (2025).
41. Green, J. et al. A cell-type-specific error-correction signal in the posterior parietal cortex. *Nature* **620**, 366–373 (2023).
42. He, Q., Liu, J. L., Eschapasse, L., Beveridge, E. H. & Brown, T. I. A comparison of reinforcement learning models of human spatial navigation. *Sci. Rep.* **12**, 13923 (2022).
43. Foster, D. J., Morris, R. G. & Dayan, P. A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus* **10**, 1–16 (2000).
44. Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A. & Uchida, N. Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron* **74**, 858–873 (2012).
45. Jeong, H. et al. Mesolimbic dopamine release conveys causal associations. *Science* **378**, eabq6740 (2022).
46. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**, 837–841 (2009).
47. Matsuda, W. et al. Single nigrostriatal dopaminergic neurons form widely spread and highly dense axonal arborizations in the neostriatum. *J. Neurosci.* **29**, 444–453 (2009).
48. Menegas, W., Babayan, B. M., Uchida, N. & Watabe-Uchida, M. Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *eLife* **6**, e21886 (2017).
49. Lerner, T. N. et al. Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* **162**, 635–647 (2015).
50. De Jong, J. W. et al. A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron* **101**, 133–151 (2019).
51. Van Elzelingen, W. et al. A unidirectional but not uniform striatal landscape of dopamine signaling for motivational stimuli. *Proc. Natl Acad. Sci. USA* **119**, e2117270119 (2022).
52. Aragona, B. J. et al. Regional specificity in the real-time development of phasic dopamine transmission patterns during acquisition of a cue-cocaine association in rats. *Eur. J. Neurosci.* **30**, 1889–1899 (2009).
53. Tsutsui-Kimura, I. et al. Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. *eLife* **9**, e62390 (2020).
54. Barter, J. W. et al. Beyond reward prediction errors: the role of dopamine in movement kinematics. *Front. Integr. Neurosci.* **9**, 39 (2015).
55. Coddington, L. T. & Dudman, J. T. Learning from action: reconsidering movement signaling in midbrain dopamine neuron activity. *Neuron* **104**, 63–77 (2019).
56. Long, C. et al. Constraints on the subsecond modulation of striatal dynamics by physiological dopamine signaling. *Nat. Neurosci.* https://doi.org/10.1038/s41593-024-01699-z (2024).
57. Markowitz, J. E. et al. Spontaneous behaviour is structured by reinforcement without explicit reward. *Nature* **614**, 108–117 (2023).
58. Lahiri, A. K. & Bevan, M. D. Dopaminergic transmission rapidly and persistently enhances excitability of D1 receptor-expressing striatal projection neurons. *Neuron* **106**, 277–290 (2020).
59. Bouabid, S. et al. Distinct spatially organized striatum-wide acetylcholine dynamics for the learning and extinction of Pavlovian associations. *Nat. Commun.* **16**, 5169 (2025).
60. Ragozzino, M. E., Mohler, E. G., Prior, M., Palencia, C. A. & Rozman, S. Acetylcholine activity in selective striatal regions supports behavioral flexibility. *Neurobiol. Learn. Mem.* **91**, 13–22 (2009).
61. Claudi, F. Mouse top detailed. Zenodo https://doi.org/10.5281/zenodo.3925997 (2020).

Methods

Animals

Male and female C57BL6 mice (14 total), aged 9–20 weeks at implantation, were used across experiments (Supplementary Table 2). The mice were housed under a 12 h–12 h light–dark cycle (light on, 21:00–09:00), and all of the experiments were performed during the dark cycle. Unimplanted mice were group-housed 2–5 per cage; after implantation, the mice were singly housed. During behavioural training and imaging, the mice were restricted to 1–2 ml of water per day so as to maintain 80–90% of initial body weight and otherwise had free access to food and water. All studies and procedures were approved by the Boston University Institutional Animal Care and Use Committee. As there were no experimental groups to be compared, there was no randomization or blinding. There were no statistical methods used to predetermine sample size.

Implant and surgeries

Multifibre arrays were fabricated in-house according to the protocol described previously[25]. In brief, 50-µm-diameter fibres (46 µm core, 4 µm cladding, 0.66 NA) were threaded through holes in a custom 3D-printed frame, cut to specific lengths targeting the striatum and fixed in place with UV-cured glue (Norland Optical Adhesive). Each implant contained 51–114 fibres (Supplementary Table 2), positioned in a pattern targeting the full AP, ML and DV extent of the striatum unilaterally (hemispheres counterbalanced across mice). A range of fibre densities was used to control for potential damage effects[25]. The non-implanted ends of the fibres were bundled and secured in a 0.8–1.3-mm-diameter polyimide tube (MicroLumen), which was cut with a razor blade and polished with fine-grained sandpaper (5 µm then 3 µm; ThorLabs). Before implantation, a calibration procedure was performed to map each implanted fibre location to the imaging surface.

Pretrained mice were fitted with a custom metal headplate[23], attached to the skull with Metabond (Parkell) several weeks before implant and injection surgeries. The headplate was secured lightly to allow later removal. After pretraining, mice received intracranial injections of AAVs carrying the genetically encoded fluorescent sensor dLight1.3b34 (pAAV-CAG-dLight1.3b, AAV9 or AAV5), diluted 1:3 in PBS (final titres: 6×10^{12}, 4.25×10^{12} or 3.75×10^{12} GC per ml). The prepared multifibre arrays were then implanted. One mouse underwent injections and implantation in two separate surgeries.

Mice were anaesthetized with isoflurane (1–3%) and positioned in a stereotaxic frame (Kopf). A large (~4.5 × 3 mm) craniotomy was drilled over the left or right striatum (Supplementary Table 2). AAV injections were targeted to 17–43 sites across the striatum to maximize expression around target fibre locations. Each site received 200–800 nl of viral solution delivered at around 100 nl min^{-1} using a pulled-glass pipette. The fibre array was mounted on the stereotaxic frame, aligned to bregma with a reference fibre, and slowly lowered into the brain. The exposed brain was sealed with Kwik-Sil (World Precision Instruments), and a thin layer of tissue glue was applied to lightly seal the medial grid edge to the skull. C&B-Metabond (Parkell) was then applied to secure the implant and seal any exposed bone or Kwik-Sil. A metal headplate, ring and plastic protective cap were affixed with Metabond, offset to the side of the implant. Finally, a layer of Metabond blackened with carbon powder (Sigma-Aldrich) was applied over the exposed fibres and protective cap.

Fibre localization

Fibre localization has been described in detail previously[25]. In brief, at the experiment end, the mice were perfused transcardially with 1% PBS followed by 4% paraformaldehyde. The implanted skulls were then prepared for CT scanning by exposing the ventral surface of the brain and soaking in a 25% Lugol's solution contrast agent. The 3D CT scans were registered to the Allen Mouse Brain Common Coordinate Framework 3D 10 µm reference atlas[62] and the fibre tips were localized and mapped to the implant array design, using a combination of FIJI/ImageJ plugins and custom MATLAB apps developed in house[25]. Atlas coordinate locations of all the fibres were generated by semi-automated fibre tip localization, and anatomical labels were assigned from the reference atlas. These labels were visually inspected and reassigned based on the boundaries in the CT image, when necessary for borderline cases. Only fibres localized to the striatum were included for analysis (Extended Data Fig. 1). All fibre locations and spatial maps are plotted relative to bregma.

Behaviour

Head-fixed behaviour and visual-cue display. To enable running in all directions, mice were headfixed over a hollow Styrofoam ball treadmill floating in a 3D-printed plastic cradle with holes to deliver pressurized air[23]. Ball rotation was measured along three axes (pitch, yaw and roll) using two optical mouse sensors (Logitech G203, hard plastic shells removed, set to 400 dpi and 1 kHz polling rate), positioned at the ball equator at 90° angles from each other. Outputs from the sensors conveying pitch and yaw rotation were converted to analogue voltages through a Raspberry Pi (3B+) at 100 Hz (maximum velocity of 3.5 m s^{-1}, corresponding to 3.3 V) and relayed to a NIDAQ board (National Instruments, PCIe-6343). The velocity direction was read as a digital binary variable from the Raspberry Pi through the NIDAQ board. For water reward delivery, a 60 ml syringe was connected through plastic tubing to a digitally gated solenoid valve (Neptune Research, 161T012), which was then connected to a post-mounted water spout. Licking was monitored with a capacitive touch circuit attached to the spout. Solenoid triggering and behavioural data acquisition were performed by the NIDAQ at 2 kHz and synchronized with custom MATLAB software.

Visual cues were designed using the Virtual Reality MATLAB Engine (ViRMEn)[63] and displayed on an array of five 12 × 21 inch monitors arranged in a semi-circle approximately 13 inches in front of the mice. Each cue consisted of vertically or horizontally oriented black-and-white stripes extending 10 inches across the entire vertical monitor surface (21 inches). For four mice, cues included two small white circles positioned symmetrically on both sides at an angle of π/5.5 rad to indicate the trial end position. Angular (yaw) ball velocity in the loc+vis instrumental task (Fig. 1) was converted into angular cue movement across the monitors in ViRMEn. ITIs without cues ranged from 7–14 s, drawn pseudorandomly from a uniform distribution. The trials began with the onset of one of the two visual cues, presented at the centre of the mouse's visual field. Each cue appeared on 50% of trials, with no more than three consecutive trials showing the same cue. Cue rotation, trial end and reward delivery were determined according to the task rules described below. After reward delivery or trial end, cues remained for 0.5 s before turning off. Task contingencies and trial structure were controlled using custom MATLAB functions.

After surgery, the mice recovered for 2–6 weeks to allow for fluorescent sensor expression. Three to five days before behavioural training, water was restricted to 1–2 ml per day to maintain body weight at 80–90% of baseline. For approximately one week before task training, mice were acclimatized to the head-fixed setup for 10–30 min per day, during which unexpected water rewards (7 µl) were delivered at random intervals drawn from a 5–30 s uniform distribution. The mice were then trained on a given task for 30–60 min per day, including 30 min of imaging time. After the imaging sessions began, training and imaging continued with no more than 1–3 day breaks, except when switching to a new task. Task type and training schedules are described below.

Loc+vis task. A total of 10 mice (7 male, 3 female; Supplementary Table 2) were trained on the loc+vis task for DA measurements (Fig. 1a). Beginning at cue onset, the cue moved laterally across the monitors at

a speed and direction determined by the ball's yaw velocity. Velocity (in rad s^{-1}) was converted into cue movement using a scaling factor of 0.02, producing trial durations with a median of 5 s. To maintain natural correspondence between directional locomotion and cue movement relative to the mouse, lateral cue movement was opposite to the mouse's running direction (that is, clockwise ball rotations corresponded to positive velocities and clockwise cue movement, experienced by the mouse as a counterclockwise turn). The trials ended when the cue rotated by π/5.5 radians (~33°) in either direction. The cue pattern (horizontal or vertical stripes) indicated which trial end position (left or right) would be rewarded, and cue-reward mapping was counterbalanced across mice.

A subset of mice included in the loc+vis analysis had been pretrained on other task variants. No differences in TE signalling or spatial patterns were observed across mice with different pretraining schedules, so all were combined for analyses. Four mice were initially trained in the vis-only task (see below). The three mice included in the learning analysis (Extended Data Fig. 10) were pretrained on a block reversal task with a separate cue before loc+vis. Mice were imaged and trained on the loc+vis task until reaching a reward rate exceeding 80% within one session, followed by an additional 4–12 days of training. Except for the learning analyses (Extended Data Fig. 10), sessions with performance below 80% were excluded from analyses.

Vis-only task. Four naive mice (2 male, 2 female) were introduced to the vis-only task (Extended Data Fig. 5a–h) before any training on other task variants. In this task, the relationships between visual cues and rewarded cue directions were identical to those in the loc+vis task, but cue movement was entirely independent of the animals' running velocity. On each trial, cue direction and speed were yoked to trials from two mice that had previously completed the loc+vis task. Congruent and incongruent cue trajectories (speed + direction) were drawn randomly from a set of correct trials in the loc+vis task with durations under 10 s, which met the inclusion criteria described in the data preprocessing section. Trials were chosen so that each cue was presented on 50% of trials, with 70% rewarded and 30% unrewarded in each session. The reward was delivered when cues reached the rewarded lateral position (same as in the loc+vis task) and omitted if the end position did not match the rewarded position. Imaging on the vis-only task was carried out for 7–8 days after animals reached the learning criterion described in the data preprocessing section.

Loc-only task. After completing the loc+vis task, two mice (1 male, 1 female) were trained and imaged on the loc-only task (Extended Data Fig. 5i–n). The task rules were identical to the loc+vis task, except that the visual cue remained stationary at the centre of the mouse's visual field and did not move laterally with ball velocity. Trial end and reward delivery were determined by calculating the theoretical position of the visual cue if it were coupled to the angular velocity as in the loc+vis task. In other words, the locomotor requirements of the mouse were identical to the loc+vis task, but no visual flow was provided to indicate the mouse's angular velocity or position relative to the goal. Mice were imaged on this task for 7–8 days, achieving performance above 80% rewarded trials.

Pavlovian task. Four mice were trained on a Pavlovian conditioning task in which a light cue was paired with a water reward (Extended Data Fig. 4c–e). During each imaging session, mice received 60 presentations of light and tone cues (30 of each) in a pseudorandom order. Each cue was presented for 6 s, with a water reward delivered at a fixed delay of 3 s from cue onset, followed by an ITI randomly drawn from a uniform distribution of 4–40 s. During each session, eight additional random rewards were delivered during ITI periods. Light cues were delivered with an LED (ThorLabs, M470L3, 470 nm) at 7 mW intensity, positioned around 20 cm from the mouse at a 45° angle contralateral to the implant. Tone cues (12 kHz, 80 dB) were delivered via a USB speaker positioned around 30 cm from the mouse. Mice were trained until they reached a learning criterion, described in the 'Analysis' section, and then trained and imaged for an additional 2–10 days. Only trials in which the light cue was delivered were included in the analysis.

Optical measurements

Fluorescence around the implanted tips of each fibre was measured by imaging the polished surface of the fibre bundle with a table-mounted microscope[25]. Excitation light was provided by high-power LEDs (470 nm and 405 nm; ThorLabs, No. SOLIS-470C, SOLIS-405C). Light power from the objective was adjusted to be within 60–75 mW, resulting in a power density of 1.4–1.7 mW mm^{-2} at the fibre tips. Emission light was band-pass filtered (Chroma, 525/50 m) and focused with a tube lens (ThorLabs, TTL165-A) onto the CMOS sensor of the camera (Hamamatsu, Orca Fusion BT Gen III). Imaging data were acquired using HCImage Live (Hamamatsu). Single-wavelength 470 nm excitation was carried out with continuous imaging at 30 Hz (33.33 ms exposure time). In 7 mice, quasi-simultaneous excitation with 405 nm and 470 nm light was conducted for 1–2 days to address potential artifactual contributions (Fig. 2d,e,i,j). Although increases in DA binding still resulted in slight (negative) fluorescence changes with 405 nm illumination, no significant TE encoding was observed (Fig. 2i,j; 0 out of 236 fibres significant for 405 nm, 103 out of 246 fibres significant for 470 nm; *t*-test on model coefficients, $P < 0.05$ for three consecutive timepoints within the 1 s cue window, Bonferroni corrected for the number of timepoints). Excitation wavelength separation was achieved by alternating the 405 nm and 470 nm LEDs and synchronizing each with imaging acquisition. Alternation was performed at 36 Hz, resulting in a frame rate of 18 Hz for each excitation wavelength (20 ms exposure time).

Analysis

All reported intervals in the text, and error bars and shaded regions in the figures indicate 95% confidence intervals. For all figures, $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$, and NS or no stars indicates $P > 0.05$. All analyses, except where otherwise noted, were performed using built-in and custom MATLAB functions and scripts.

Behavioural variables and session inclusion criteria. Angular velocity was calculated from ball yaw velocity, converted to rad s^{-1}, and signed according to one of three conventions, as indicated in each plot: (1) by direction (+, left; −, right); (2) by direction relative to the implant (implant-signed; +, contralateral; −, ipsilateral), and (3) by congruence (congruence-signed; +, congruent; −, incongruent). Linear velocity was defined based on the pitch and roll ball velocities, converted into m s^{-1} and combined so that

$$v_{\mathrm{lin}} = (v_{\mathrm{pitch}}^2 + v_{\mathrm{roll}}^2)^{-2}$$

Linear and angular velocities were smoothed using a LOWESS filter with a 0.3 s smoothing window, then low-pass filtered with a 1.5 Hz cut-off. Linear acceleration was calculated as the derivative of linear velocity. Ipsilateral and contralateral acceleration were computed as the derivative of the absolute value of angular velocity, separated into periods when the mouse's velocity direction was either ipsilateral or contralateral to the implant (Extended Data Fig. 4a,b). Angular acceleration was calculated as the derivative of implant-signed angular velocity (+, contralateral; −, ipsilateral), such that positive values indicate increasing velocity in the contralateral direction or decreasing velocity in the ipsilateral direction, and negative values indicate decreasing velocity in the contralateral direction or increasing velocity in the ipsilateral direction. Linear and angular acceleration were smoothed with a LOWESS filter (0.3 s window) and low-pass filtered with a 1.5 Hz cut-off.

Congruence was defined based on whether yaw velocity was directed toward (congruent) or away from (incongruent) the rewarded (goal) position. TE was defined as the magnitude of angular velocity (rad s^{-1}), signed according to congruence (+, congruent; −, incongruent). For the loc+vis and loc-only tasks, TE and congruence were calculated as the average angular ball velocity within a 0.3 s window preceding cue onset, chosen to occur before the average cue-evoked velocity change. For the vis-only task, cue-based TE and congruence on each trial were calculated as the average angular velocity of the cue within a 0.2 s window following cue onset.

For the vis-only task, a lick index was calculated to assess learning of the cue-direction associations (Extended Data Fig. 5a–c) by comparing the sum of the licking 1 s before reward (L_{rew}) and 1 s before cue onset ($L_{baseline}$): $(L_{rew} - L_{baseline})/(L_{rew} + L_{baseline})$. Days were included if they met two criteria: (1) the average lick index for rewarded trials was significantly greater than zero, and (2) the average lick index for rewarded trials was significantly greater than that for unrewarded trials (6–7 days, $P < 0.01$, one-sided shuffle test with 10,000 iterations). For criterion 1, the average lick index was compared to an empirical null distribution generated by randomizing the labels for reward and baseline time periods. For criterion 2, the difference in lick index between rewarded and unrewarded trials was compared to an empirical null distribution generated by randomizing the trial-type labels and recomputing the lick difference 10,000 times. For the Pavlovian task, the lick index was computed as $(L_{rew} - L_{baseline})/(L_{rew} + L_{baseline}) \times L_{rew}$, and normalized to the maximum single-trial lick index across mice.

To quantify running direction biases for each mouse, we computed a contralateral bias index based on the time to reward for each cue (Extended Data Fig. 6g):

$$B_{contra} = (t_{ipsi} - t_{contra})/(t_{ipsi} + t_{contra})$$

where t_{ipsi} and t_{contra} are the natural logarithm of the time from cue onset to reward, averaged across ipsilateral and contralateral trials, respectively. The majority of mice were biased to run in the direction contralateral to the implant before cue onset, resulting in faster rewards and higher reward rates for the cue associated with the contralateral direction (Extended Data Fig. 6j–l). This bias could have arisen from the offset positioning of the headplate toward the implanted hemisphere, variability in intrinsic biases across mice or potential brain damage from the implants. We used implants with a range of fibre densities (Supplementary Table 2) to control for possible brain damage and observed no correlation between the bias index and the number of implanted fibres ($\rho = -0.33$, $P = 0.35$, $n = 10$ mice, Pearson's correlation coefficient). A subset of mice that rarely ran in the ipsilateral direction before the cue onset was excluded from certain analyses (Extended Data Fig. 3g,h (middle)).

Except for the learning analysis (Extended Data Fig. 10), only sessions in the loc+vis and loc-only tasks in which performance exceeded 80% correct were included (3–10 days). Data were concatenated across sessions. Trials were excluded if the total velocity—computed as the square root of the sum of squares of ball velocity (in m s^{-1}) across all three axes (pitch, yaw, roll) and smoothed as described above for linear velocity—dropped below 0.3 m s^{-1} within 0.5 s before cue onset, indicating that the mouse was at rest. This threshold was selected manually based on the distribution of pre-cue total velocities. In the vis-only task, sessions in which the lick index was below 0.15 were excluded.

Signal pre-processing, Δ*F*/*F* extraction and fibre inclusion criteria. Videos of the fibre bundle surface were processed to extract changes in fluorescence from each fibre, as described in detail previously[25]. Videos were first motion-corrected using a cross-correlation algorithm described previously[25]. Mean fluorescence time series were extracted from circular regions of interest (around 25 μm diameter) corresponding to each fibre. Relative changes in fluorescence (Δ*F*/*F*) were calculated by dividing the mean fluorescence time series by the eighth percentile fluorescence over a 30 s sliding window, then subtracting the median across the recording session.

Fibre inclusion criteria. Fibres were included in the analysis if they were conclusively determined to be in the striatum (see the 'Fibre localization' section above) and if they had an estimated signal-to-noise ratio (SNR) greater than 1 (Extended Data Fig. 2). To estimate the SNR, we used an unbiased approach based on two assumptions about spectral densities of the dLight signal: (1) white noise due to motion or noise introduced during image collection has a flat power spectrum across frequencies; and (2) the dLight1.3b signal is limited to lower frequency bands (<10 Hz) due to sensor kinetics. To estimate the total noise power, we calculated the average spectral density within the 10–14 Hz frequency band and multiplied this by the full frequency range (14 Hz). To estimate the total signal power, we calculated the total power in the full frequency range (0–14 Hz) and subtracted the noise power. The SNR was then computed by dividing the signal power by the total noise power. Power estimates were computed using the bandpower function in MATLAB and divided by the bandpass range.

Temporal dynamics. To assess the timescales of DA signalling (Fig. 1f–i), we constructed autocorrelograms for each recording location by computing Pearson's correlation coefficients using the corrcoef function in MATLAB on the full Δ*F*/*F* trace concatenated across days with >75% correct performance, with time shifts up to 2 s. The resulting autocorrelograms were normalized to the maximum correlation coefficient. To determine the decay constant (Fig. 1h), we fit each autocorrelogram to an exponential model using the fit function in MATLAB:

$$y = ae^{bt} + ce^{dt}$$

where y is the autocorrelogram trace and t is the time shift in seconds. The decay constant was determined as the minimum of b and d, indicating the faster decay term. The second, slower decay term accounts for noise, expressed as positive correlations at longer time lags. Autocorrelograms in Fig. 1g have the noise term subtracted out.

Mean comparisons and LME models. Comparisons of two measures across recording sites and mice (Fig. 2e, j and Extended Data Figs. 2d, 3c,f,h, 5g,h,n and 6c,f,i) were performed using a two-sided t-test on the main effect of measure type in a LME model, using the MATLAB function fitglme, with random effects of mouse identity on the intercept and measure type to account for mouse-to-mouse variability, and a random effect of recording site identity on the intercept to enable paired comparisons.

Averages of coefficients or Δ*F*/*F* (Figs. 2d,i and 4a and Extended Data Figs. 3b,e,g, 5g,h,m and 6a,b,d,e,g,h) were obtained from the intercept of an LME with a random effect of mouse identity on the intercept to account for mouse-to-mouse variation, using the MATLAB function fitglme. 95% confidence intervals (shaded regions) were estimated using the MATLAB function coefCI on the LME outputs. A separate LME was fit for each timepoint.

Regression models. To quantify encoding of TEs and cue value at each recording site (Fig. 2h–j and Extended Data Fig. 6), Δ*F*/*F* traces (concatenated across included sessions) were modelled using linear regression with sets of FIR functions. Each FIR (s_t), equals 1 at timepoint t relative to cue onset and 0 at all other timepoints across a 2 s post-cue onset window. Thus, a total of 2 s × frame rate (Hz) FIR functions tiled the window after cue onset. To estimate the time course of TE, cue identity (CI), reward history (RH) and upcoming reward (UR) encoding, each set of FIR functions was multiplied by a variable, x, which was constant across each 2 s cue window, but varied across trials according to the corresponding trial level variable. Four continuous locomotor variables

(a^i, linear velocity, linear acceleration, angular velocity and angular acceleration) were also included as covariates. The following linear function was fit to estimate coefficients β using MATLAB's fitlm function:

$$y(t) = \sum_{t=0}^{2s} s_t(\beta_t^0 + \beta_t^{TE}x^{TE} + \beta_t^{CI}x^{CI} + \beta_t^{RH}x^{RH} + \beta_t^{UR}x^{UR}) + \sum_{i=1}^{4} \beta^i a^i$$

where $y(t)$ is $\Delta F/F$, β_t values are coefficient estimates at each timepoint t from the cue onset, and β^i are coefficients for the locomotion variables. Trial level variable definitions are as follows:

x^{TE} (trajectory error): scaled by average angular velocity in the 0.3 s before cue onset, with sign determined by congruence of movement direction. β_t^{TE} represents the TE coefficients reported throughout the paper.

x^{CI} (cue identity): 1 for contralateral cues, −1 for ipsilateral

x^{RH} (reward history): 0 if the previous trial was rewarded, −1 if unrewarded.

x^{UR} (upcoming reward): 0 on rewarded trials and −1 on unrewarded trials.

The intercept term, β_t^0, represents the cue-evoked DA response after accounting for all other variables, and is referred to as 'non-TE' throughout. As β_t^0 reflects the response when all regressors are zero, we defined '0' as rewarded trials (and trials following rewarded trials), ensuring that it represents cue responses during high-performance conditions. RH and UR were coded with −1 (rather than +1) on unrewarded trials so that their coefficients could be directly interpreted as $\Delta F/F$ differences between rewarded and unrewarded conditions. Similarly, +1 and −1 were used for contralateral and ipsilateral cues, respectively.

We tested contributions of each variable to TE encoding by refitting the model with each set of regressors removed (Extended Data Fig. 3a–c). To compare continuous TE encoding with binary congruence encoding, we replaced x^{TE} with x^{cong} (1, congruent; −1, incongruent; Extended Data Fig. 3d–f). As these regressors have different scales, comparisons were made on t-statistics rather than coefficient magnitudes.

To compare TE encoding across trial types (for example, congruent versus incongruent, contra versus ipsi cue, contra versus ipsi angular velocity), we included separate TE FIR functions for each condition (Extended Data Fig. 3g,h). For contra versus ipsi initial velocity analyses, three mice were excluded due to directional bias (<30 trials in one direction). For the vis-only task (Extended Data Fig. 5a–h), two TE terms were modelled: a cue TE, defined as the mean visual cue velocity (direction and magnitude) relative to the rewarded direction, averaged in a 0.2 s window after cue onset, and a locomotion TE, defined as in the loc+vis task. To compare the loc-only and loc+vis tasks, loc+vis data were refit with cue TE terms defined identically to the vis-only task. For the Pavlovian task (Extended Data Fig. 4), TE terms were replaced with trial-by-trial angular velocity magnitude (0.3 s pre-cue).

The significance of model coefficients (Fig. 4d,g and Extended Data Fig. 7a) was determined using a t-test on the FIR coefficients within a 1 s cue window. A fibre was determined significant if ≥3 consecutive timepoints met $P < 0.05$ after Bonferroni correction for the number of timepoints within the 1 s window.

Spatial gradient analysis. Spatial coefficients and gradient vectors. We estimated spatial coefficients for each variable by fitting a linear mixed-effects model to the maximum FIR coefficient after cue onset as a function of fibre position (AP, ML, DV, relative to bregma):

$$\mathbf{y} = \beta^0 + \beta^{ML}\mathbf{x}^{ML} + \beta^{AP}\mathbf{x}^{AP} + \beta^{DV}\mathbf{x}^{DV} + \text{random effects}$$

$$= \beta^0 + \sum_{d=[AP,ML,DV]} (\beta^d \mathbf{x}^d) \text{random effects}$$

where $\mathbf{y}$ is a $1 \times n$ vector of the maximum coefficient t-statistics for each of n fibres. $\mathbf{x}^d$ is the fibre position in each axis. Random effects of mouse identity were included on the intercept and position terms in each dimension to account for mouse-to-mouse variability. The total gradient vector was defined as $\mathbf{v} = [\beta^{ML}, \beta^{AP}, \beta^{DV}]$, representing the direction and magnitude of maximal variation in $\mathbf{y}$.

To assess consistency across mice, we computed gradient vectors for each mouse ($\mathbf{v}_m$) using fixed-effects regression. We then tested for directional concentration with a weighted Rayleigh's test, weighting by vector length[64]. The resultant vector $\mathbf{R}$ was

$$\mathbf{R} = \sum_{m=1}^{n} [\beta^{ML}, \beta^{AP}, \beta^{DV}]_m = \sum_{m=1}^{n} \mathbf{v}_m$$

where n is the number of mice. The weighted mean resultant vector length (Rayleigh's coefficient, $\bar{R}$) was

$$\bar{R} = \left\| \mathbf{R} / \sum_{m=1}^{n} \left\| \mathbf{v}_m \right\| \right\|,$$

The Rayleigh's test statistic was defined as

$$S = p \times n \times \bar{R}^2$$

where $p = 3$ is the number of dimensions, and n is the number of mice. S was corrected for small sample sizes:

$$S^* = (1 - 2/2n)S + (2n(p+2))^{-1}S^2.$$

S^* was compared against a null distribution that assumes vector directions that are uniformly distributed across the sphere. Under the null hypothesis, S^* should follow a χ^2 distribution with p degrees of freedom, so we compared S^* against χ_3^2 using the chi2cdf function in MATLAB. Significance was set at $P < 0.05$, Bonferroni corrected where noted.

Comparing spatial coefficients across variables. To test whether spatial coefficients differed between two variables, we fit a linear mixed effects model including location, variable type and an interaction between location and variable type (Figs. 4j and 5j):

$$[y_1; y_2] = \sum_{d=[AP,ML,DV]} (\beta^0 + \beta^d[x^d; x^d] + \beta^{ds}s[x^d; x^d] + \beta^s s + \text{random effects})$$

where β^{ds} is the interaction coefficient, y_1 and y_2 are the two sets of coefficients, and d is the spatial dimension.

Interaction coefficients β^{ds} indicating differences between the coefficients were considered significant when $P < 0.05$, Bonferroni corrected for multiple comparisons, where noted. To test for differences across the cue-value variables (Fig. 4h), we repeated this procedure for three groups and performed an ANOVA on the interaction term.

Comparing gradient vectors across variables. We compared mean gradient directions between variables using the high-concentration F-test, which evaluates whether two sets of unit vectors differ in orientation[64] (Figs. 4k and 5k). For groups with resultant lengths R_1, R_2 and combined R, the test statistic was:

$$[(R_1 + R_2 - R)/(p-1)][(n - R_1 - R_2)/(n-2)(p-1)] \sim F_{p-1,(p-1)(n-2)}$$

where n is the number of mice and p is the number of dimensions. This test assumes concentrated samples ($\bar{R} > 0.67$). For vectors with significantly different mean directions, we estimated the mean angle between the two vectors (Figs. 4l and 5k) with a cosine similarity index:

$$C = (\mathbf{v}_{1m}/||\mathbf{v}_{1m}||) \cdot (\mathbf{v}_{2m}/||\mathbf{v}_{2m}||)$$

averaged across mice and weighted by the product of vector lengths where $\mathbf{v}_{1m}$ and $\mathbf{v}_{2m}$ are the vectors from each sample in mouse m. We then

took an average across mice, weighted by the product of the lengths of the two vectors $||\mathbf{v}_{1m}|| \times ||\mathbf{v}_{2m}||$. The mean angle was obtained as:

$$\theta = (180/\pi)\cos^{-1}(C)$$

where θ is the average angle between the two vectors, in degrees (Fig. 4l). 95% confidence intervals were constructed using a bootstrap distribution in which we resampled the cosine similarity index 100,000 times.

For comparisons across three groups (Fig. 4i), we extended the high-concentration *F*-test to three-way designs and compared the statistic to $F_{2(p-1),(p-1)(n-3)}$.

Acceleration correlations in the ITI. To determine the relationship between DA release and acceleration independent of the task, we modelled $\Delta F/F$ traces during the ITI periods, starting 1 s after reward delivery until the next cue onset, as a function of linear and angular acceleration at different time shifts. The $\Delta F/F$ trace was first high-pass filtered with a 0.8 Hz threshold to remove slow fluctuations that could result from haemodynamic artifacts. A separate model was fit for frame-by-frame temporal offsets between $\Delta F/F$ and ipsilateral and contralateral acceleration from −0.5 to 0.5 s. At each time lag, the following model was fit:

$$y = \beta^0 + \beta^{\text{lin}}x^{\text{lin}} + \beta^{\text{contra}}x^{\text{contra}} + \beta^{\text{ipsi}}x^{\text{ipsi}}$$

where y is the high-pass-filtered $\Delta F/F$, x^{lin} is linear acceleration, x^{contra} is the contralateral acceleration during time periods when the angular velocity is in the direction contralateral to the implant, 0 otherwise, and x^{ipsi} is ipsilateral acceleration during time periods when the angular velocity is in the direction ipsilateral to the implant, 0 otherwise. x^{ipsi} and x^{contra} are positive when the angular velocity is increasing and negative when the angular velocity is decreasing in either direction. This is in contrast to the implant-signed angular velocity described previously. A subset of mice that only ran in one direction during the ITI were excluded from this analysis.

Learning analysis. To determine how the cue response evolves across learning (Extended Data Fig. 6b), $\Delta F/F$ aligned to cue onset for each site were smoothed across trials using a moving average window of 100 trials, and baseline corrected to the first 50 trials.

To determine how DA TEs evolve across learning, we performed a dynamic regression analysis adapted from the PsyTrack approach[35]. In brief, the magnitude of the DA cue response during a 0.2 s window around the time of maximal TE encoding was modelled as linear-Gaussian model of the TE x_t,

$$y_t \sim N(\beta_t x_t, \sigma_y^2)$$

where the TE coefficients, β_t, change on each trial based on a Gaussian random walk,

$$\beta_t \sim N(\beta_{t-1}, \sigma_\beta^2)$$

The smoothness hyperparameter, σ^2_β, governs the rate of change of the TE coefficients as described previously[35]. We fit this model separately to data from each fibre. The time course of β_t was inferred from the length R sequence of DA magnitudes and TEs, where R is the number of trials across learning. β_t, σ^2_y, σ^2_β was learned for each fibre using maximum a posteriori (MAP) estimation. The smoothness hyperparameter was determined by grid search: the MAP estimate of β_t, σ^2_y was computed for ten values of σ^2_β in a fixed range (10^{-4} and 10^{-1}) and the model with the highest test log-likelihood was selected. Five-fold cross-validation was performed to find the optimal model for each fibre. The MAP estimate was determined using a quasi-Newton method (BFGS) from the Python scipy library. Further details of the PsyTrack dynamic regression approach can be found in a previous publication[35].

The dynamic TE coefficient was considered to be significant when the lower bound of the 99% confidence interval surpasses 0. Performance was considered to be significant when the reward rate, computed with a 50 trial sliding window exceeds the 99% confidence interval of an empirical null distribution computed by shuffling the trial number labels and recomputing the sliding window reward rate 10,000 times.

NNMF analysis. To decompose the learning time course of TE encoding across recording sites for each mouse into an early and late component, we used non-negative matrix factorization (NNMF). NNMF minimizes the root mean squared residual D between the matrix of TE coefficients, A^{tn} and $W^{nk} \times H^{kt}$. A is the matrix of TE coefficients across recording sites, n, and trials, t. We used $k = 2$, to decompose A into a linear combination of 2 factors, $\mathbf{h}_1$ and $\mathbf{h}_2$, which are vectors of size t (Extended Data Fig. 10c). We used NNMF instead of other linear decomposition methods like principal component analysis for interpretability because the TE weights are always positive. Thus, the time course a_i at fibre i can be modelled as a linear combination of the two components: $a_i = w_{i1}\mathbf{h}_1 + w_{i2}\mathbf{h}_2$, where $w_{ij} > 0$ are strictly positive. As NNMF uses an iterative process to identify local minima, and therefore depends on the initialization of W and H, we initialized $\mathbf{h}_1$ and $\mathbf{h}_2$ as two Gaussian impulses with a s.d. of 100, centred around an early and late trial, respectively. For mouse 1 and 3, the Gaussian impulses had means at trials 100 and 400, and for mouse 2 the Gaussian impulses had means at trials 100 and 1,000. We performed NNMF on these inputs using MATLAB's nnmf function. We labelled the output components, $\mathbf{h}_{\text{early}}$ and $\mathbf{h}_{\text{early}}$ of H^{2t} as early and late through visual inspection (Extended Data Fig. 10c), and then performed the spatial analysis on the difference in the weights, $w_{\text{late}} - w_{\text{early}}$, across sites (Extended Data Fig. 10e–g).

TD RPE model simulation. A reinforcement learning agent was trained on a stimulus response task similar to the task from Fig. 1. At each time step, the agent selects one of two actions: accelerate left or accelerate right. Based on the selected action, the angular velocity (ω) at each time step is determined by the following equations:

$$\text{If action = left: } \omega(t + \Delta t) = \omega(t) + 0.05(-1 - \omega(t))$$

$$\text{If action = right: } \omega(t + \Delta t) = \omega(t) + 0.05(+1 - \omega(t))$$

Thus, whether and how much the velocity changes is determined by the current velocity, and the direction of the change is determined by the selected action. Here $v < 0$ indicates right velocities and $v > 0$ indicates left velocities. These equations bound the velocity by +1 and −1, and the maximum acceleration at $0.05/\Delta t$.

The state space input used for action selection and learning (s_0) includes 10 units each representing evenly tiled Gaussian tuning curves, for both the continuous position and velocity variables (separate units for position and velocity). Three units represent the binary cue identity, or goal, as a categorical variable: one unit for the left cue, one unit for the right cue and one unit for no cue. These 23 variables are provided as inputs with random weights to a second layer of $N = 1{,}000$ neurons, according to the following equations:

$$s = \text{ReLU}(W_s s_0)$$

where W_s is a $N \times 23$ encoder matrix drawn randomly from a standard normal distribution and ReLU is the rectified linear function, which maps negative inputs to zero and otherwise leaves them unchanged.

At each timestep, the action is selected according to probabilities given by a softmax function of the actor output:

$$a = W_a s$$

$$P(a) = e^a / \Sigma e^a$$

where W_α is a $2 \times N$ matrix representing the actor weights, updated at each timestep with a learning rate of $\alpha = 0.25/N$ using the following equation:

$$W_\alpha(t + \Delta t) = W_\alpha(t) + \alpha\delta(t)a(t)[s(t)]^T$$

$\delta(t)$ represents the temporal difference reward prediction (tdRPE), which models DA release and is computed as follows:

$$\delta(t+1) = r(t+1) + \gamma v(t+1) - v(t)$$

where r indicates the presence (1) or absence (0) of reward, $\gamma = 0.9$ sets the temporal discount factor and $v(t)$ is the value function determined by the following equations:

$$v(t) = \mathbf{w}_v \cdot s(t)$$

$$\mathbf{w}_v(t+1) = \mathbf{w}_v + \alpha\delta s(t)$$

$\mathbf{w}_v$ is a vector of critic weights used to estimate value. On each trial, the position (x) is initialized at $x_0 = 0$, and the velocity (ω) is initialized randomly based on a normal distribution centred around $\mu = \pm 0.25$ (sign determined at random for each simulation, but consistent across trials) to mimic the contralateral bias observed in our mice. For unbiased simulations (Extended Data Fig. 7d–f), initial velocity distribution centred around $\mu = 0$. The position is updated on each timestep based on the velocity, according to $x(t+1) = x(t) + \omega(t)$. One of two visual cues (left or right) is presented randomly with 50% probability at timestep 5, and the trial finishes when the agent reaches the end position $x = \pm 5$. If the left cue is present, the agent will be rewarded for reaching $x = +5$, for the right cue, the agent will be rewarded for reaching $x = -5$. We ran 10 simulations of 10,000 trials and analysed the RPE, δ_t across trials and simulations (Fig. 3a–j).

To test the effect of the state space s, on the RPE δ, we repeated this procedure with modified state spaces (Fig. 3k–n and Extended Data Fig. 7g–l). We first tested state spaces without the position and or velocity inputs to the second mixing layer (no position, no velocity, cue only). We then tested state spaces with all inputs or only the cue input, but without the mixing layer (no mix, and cue only, no mix). To quantify the RPE cue increase and encoding of TE, reward history and cue identity, we performed a linear regression to model the RPE averaged across time $t = 1{,}0$, $\delta_{t=0,1}$ as a linear function of each variables:

$$\delta_{t=0,1} = \beta^0 + \beta^{\mathrm{TE}}x^{\mathrm{TE}} + \beta^{\mathrm{CI}}x^{\mathrm{CI}} + \beta^{\mathrm{RH}}x^{\mathrm{RH}}$$

As in our DA regression model, the cue response was estimated from β^0. Estimated coefficients were averaged across simulations for each state space.

Reporting summary

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

Data availability

All processed fluorescence, behavioural and fibre localization data have been deposited at Zenodo[65] (https://doi.org/10.5281/zenodo.17653000). Additional data are available on reasonable request from the corresponding author.

Code availability

All code used in this study for key analyses was written in MATLAB or Python and is available at Zenodo[65] (https://doi.org/10.5281/zenodo.17653000).

62. Wang, Q. et al. The Mouse Brain Common Coordinate Framework: a 3D reference atlas. *Cell* **181**, 936–953 (2020).
63. Aronov, D. & Tank, D. W. Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. *Neuron* **84**, 442–456 (2014).
64. Mardia, K. V. & Jupp, P. E. *Directional Statistics* (Wiley, 1999); https://doi.org/10.1002/9780470316979.
65. Brown, E. H. et al. Data and code for 'Striatum-wide dopamine encodes trajectory errors separated from value'. *Zenodo* https://doi.org/10.5281/zenodo.17653000 (2026).

Acknowledgements This work was supported by a Klingenstein-Simons's Foundation fellowship, Whitehall Foundation Fellowship, National Institute of Mental Health (R01 MH125835) to M.W.H.; NIMH F31NS127536-01A1 to E.H.B.; NIMH F32MH120894 to M.-A.V.; and JSTPN Early Stage Training in Neuroscience Award to Y.Z. We thank the staff at Boston University Centers for Neurophotonics and Systems Neuroscience for financial and technical support; the members of the Micro CT core, especially S. Holder, for providing equipment and technical expertise for micro-CT scanning; the staff at the Boston University Animal Science Center for providing central laboratory and animal care and support resources; L. Tian and Y. Li for providing dLight1.3b and Ach3.0 viral vectors; and M. Valyear, C. Stern and M. Hasselmo for feedback on a draft of the manuscript.

Author contributions E.H.B. and M.W.H. conceived the project, designed experiments, interpreted the data and wrote the manuscript. E.H.B. performed experiments and conducted formal statistical analyses. Y.Z. and B.D. developed and wrote code for the learning analyses, with input from E.H.B. and M.W.H.; M.-A.V., S.B., E.H.B. and M.W.H. contributed to development and validation of the microfibre array method. J.L. and A.L.-K. conceived and wrote initial code for the tdRL model approach with contributions from E.H.B.; A.A. and C.G.-N. contributed to mouse behaviour training and fibre array construction.

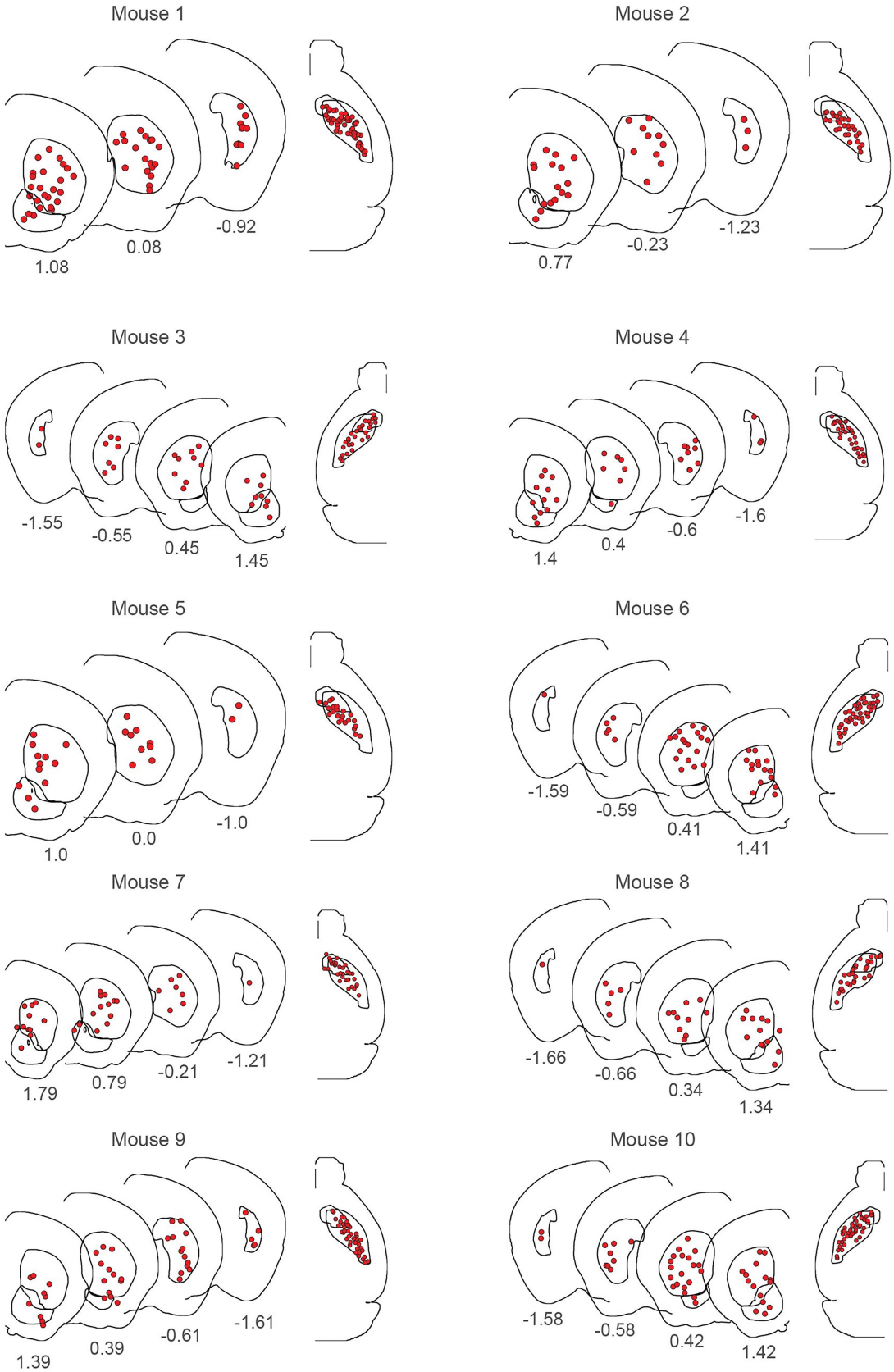
Competing interests The authors declare no competing interests.

Additional information

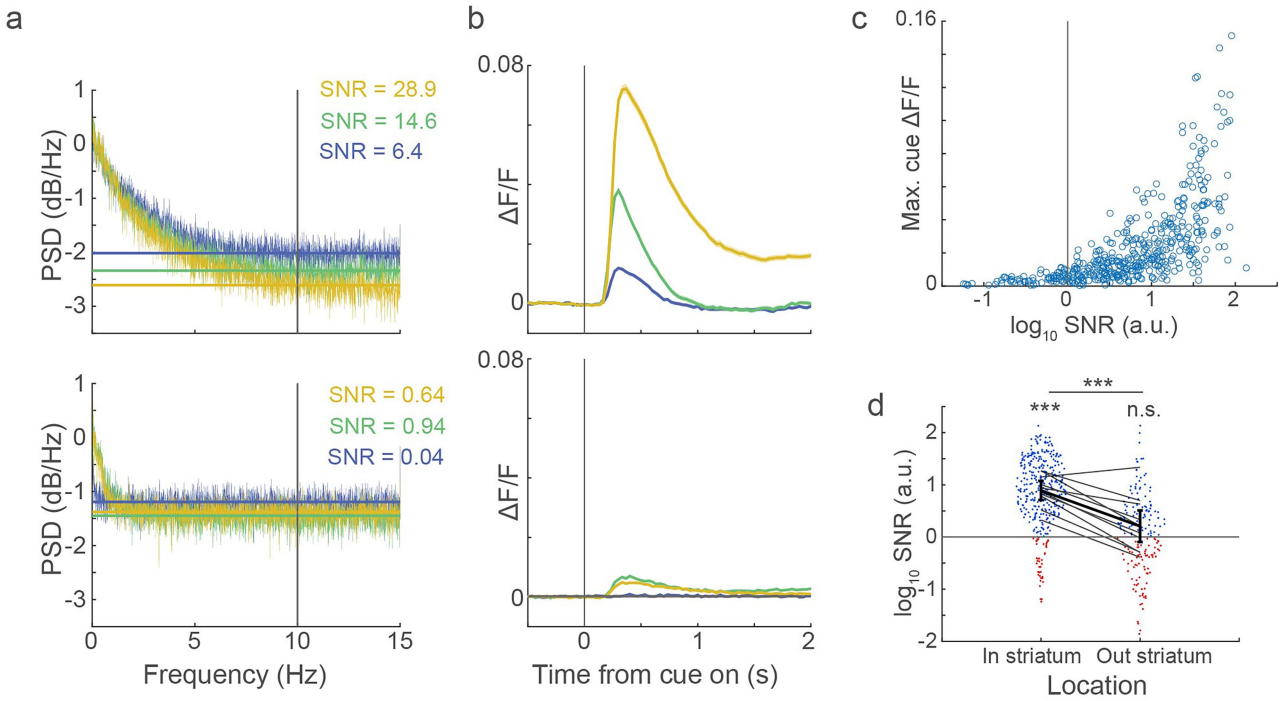
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-10083-1.

Correspondence and requests for materials should be addressed to Mark W. Howe.

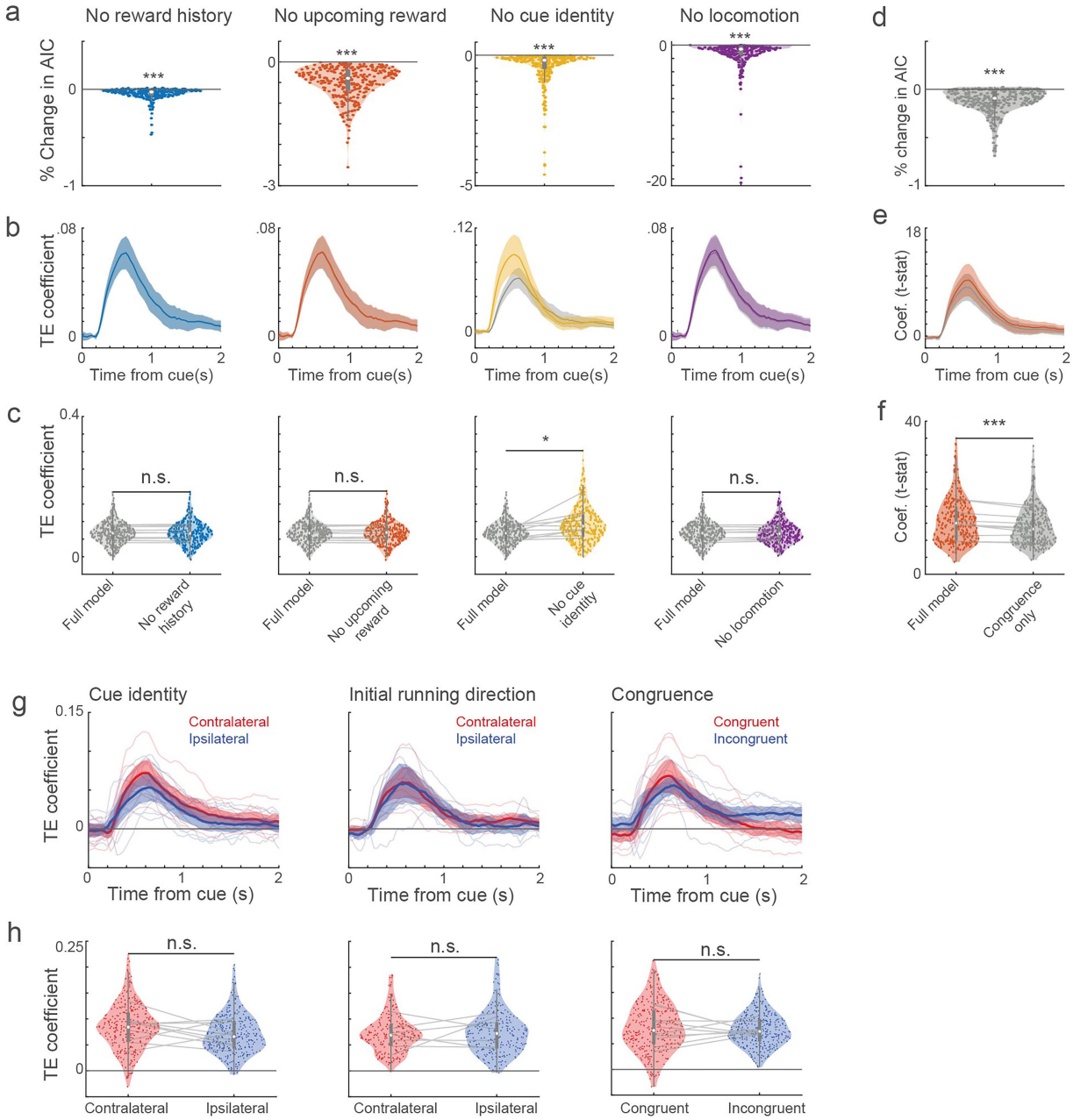
Peer review information *Nature* thanks Patricia Janak, Alexxai Kravitz and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



Extended Data Fig. 1 | Locations of all fibers across individual mice. Atlas registered locations of all fibers (red circles) included in analysis for each of 10 mice plotted in the coronal and horizontal planes relative to atlas common coordinate framework (see Methods). Each coronal plane represents a slice with a 1 mm thickness. Numbers beneath each coronal section are AP coordinates of the most anterior section of the slice, in mm relative to bregma. A few fibers appear outside of the striatal boundaries (i.e. in mice 7, 8 and 9), due to slight deviations in the actual striatum boundaries of individual mice, as determined by the CT image, from the atlas striatum boundaries. Brain schematics were adapted from the Allen Mouse Brain Common Coordinate Framework (CCFv3) (https://atlas.brain-map.org/).

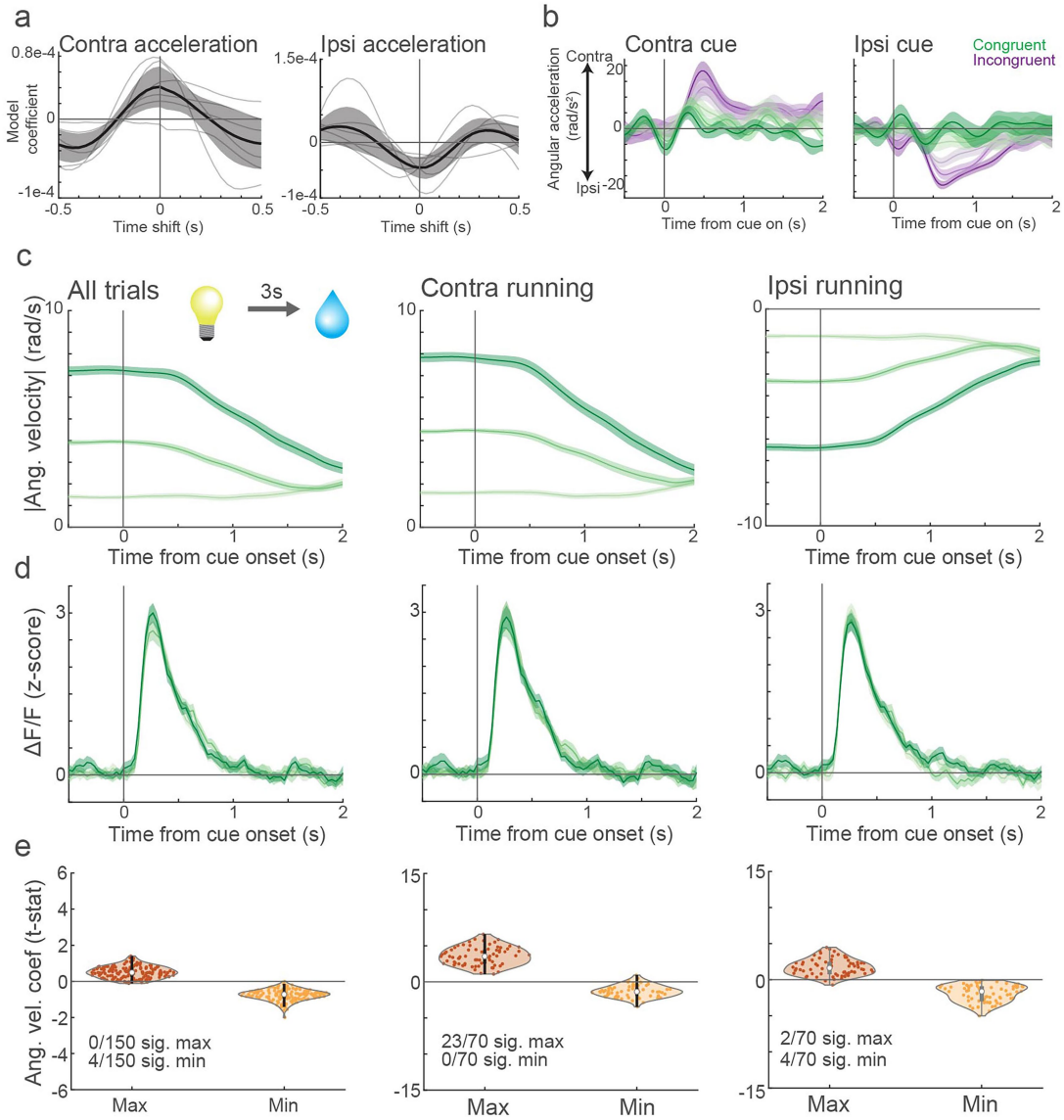


Extended Data Fig. 2 | Signal-to-noise ratios for determining recording site inclusion. a, Power spectral density (PSD) plots for example sites with high (included sites, top) and low (excluded sites, bottom) signal-to-noise ratios (SNRs). Horizontal line indicates the noise power density, estimated as the average power density within the 10-14 Hz frequency band (see Methods). **b**, Average baseline-corrected ΔF/F aligned to cue onset on all congruent trials for the same example fibers in a. **c**, Peak cue-evoked ΔF/F as a function of log SNR, each data point represents one recording site, n = 638 sites across 10 mice. The vertical line represents the point where the signal is equal to the noise (SNR = 1). Recording sites to the left of the vertical line (SNR < 1) were excluded from future analyses. **d**, Violin plots showing the log-transformed SNR for sites classified (via micro-CT localization) as within (n = 356 sites) or outside of (n = 182 sites) the striatum. SNR > 1 in blue, SNR < 1 in red. Thin lines represent average SNR for individual mice. Error bar shows mean and 95% confidence interval estimated from linear mixed effects models (in striatum: $t_{355} = 9.60$ $p = 4.33 \times 10^{-19}$, n = 356 sites, 10 mice; out striatum: $t_{181} = 1.35$, $p = 0.53$, n = 182 sites, 10 mice, two-sided t-test on model intercepts). Effect of recording site classification on SNR estimated from a linear mixed effects model, two-sided t-test $t_{636} = 8.62$, $p = 1.62 \times 10^{-16}$, n = 638 sites, 10 mice. P-values were *post hoc* Bonferroni corrected for three comparisons.

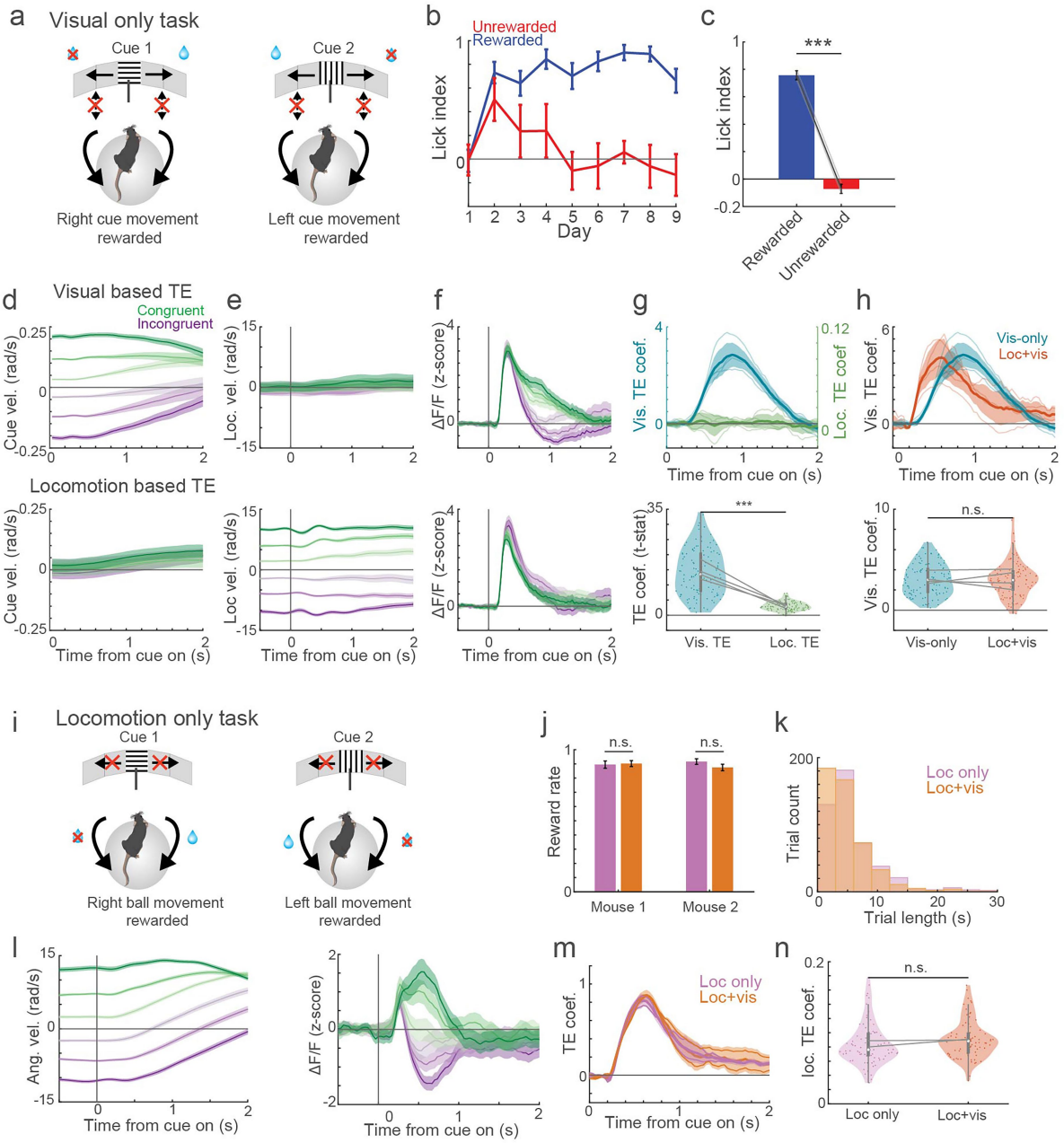


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Trajectory error encoding is present across trial types and is not explained by encoding of other task variables. a, Violin plots of the percent change in the akaike information criterion (AIC) when each variable is removed from the model (removed variable indicated above each plot). Negative values indicate worse model fit when the variable is removed (see Methods for variable definitions). Reward history: $t_{630} = 4.53$, $p = 2.85 \times 10^{-5}$; upcoming reward: $t_{630} = 6.64$, $p = 2.71 \times 10^{-10}$; cue identity: $t_{630} = 3.80$, $p = 6.36 \times 10^{-4}$; locomotion: $t_{630} = 5.42$, $p = 3.45 \times 10^{-7}$, two-sided t-test on linear mixed effects model for main effect of model type on AIC, p-values bonferroni corrected for 4 comparisons ($n = 316$ sites, 10 mice). Each data point represents one site. **b**, TE coefficients from models with (grey) and without (coloured) each variable, averaged across sites and mice using a linear mixed effects model ($df = 315$, $n = 316$ sites, 10 mice). **c**, Violin plots of the peak TE coefficient magnitudes across recording sites in models with (grey) and without (coloured) each variable. TE coefficients were unaffected by exclusion of each variable (except cue identity). Reward history: $t_{630} = 0.74$; $p = 1.0$; upcoming reward: $t_{630} = -0.65$; $p = 1.0$; cue identity: $t_{630} = -2.69$; $p = 0.30$; locomotion: $t_{630} = -1.02$; $p = 1.0$; two-sided t-test on main effect of model type on TE coefficient magnitude in a linear mixed effects model (n = 316 sites, 10 mice), bonferroni corrected for 4 comparisons. **d**, Percent change in AIC comparing the full model to a model with the TE term replaced with a binary congruence only term (no scaling with angular speed), presented as in **a**. $t_{630} = 4.52$, $p = 7.47 \times 10^{-6}$, $n = 316$ sites, 10 mice, two-sided t-test on effect of model type in a linear mixed effects model. **e**, Model coefficient t-statistics for TE (orange) and congruence only (grey) averaged across recording sides and mice using a linear mixed effects model ($df = 315$, $n = 316$ sites, 10 mice). **f**, Violin plot showing the maximum coefficient t-statistics for TE and congruence coefficients in the full model and congruence only model respectively across recording sites and mice. $t_{630} = -4.24$, $p = 2.59 \times 10^{-5}$, $n = 316$ sites, 10 mice, two-sided t-test on effect of model type on coefficient t-stat in a linear mixed effects model. **g**, TE coefficients averaged across sites and mice using a linear mixed effects model for trials split by cue identity (left, $df = 315$, $n = 316$ sites, 10 mice), initial running direction (middle, $df = 233$, $n = 234$ sites, 7 mice, only mice with >30 trials in each direction were included), and congruence (right, $df = 315$, $n = 316$ sites, 10 mice) independently. **h**, Violin plots comparing TE coefficients computed on trials split by cue identity (left), initial rotational velocity direction (middle), and congruence (right). Cue identity: $t_{630} = -1.23$, $p = 0.66$, n = 316 sites, 10 mice; rotational velocity direction: $t_{466} = -0.65$, $p = 1.00$, $n = 234$ sites, 7 mice; congruence: $t_{630} = -0.24$, $p = 1.00$, $n = 316$ sites, 10 mice, two-sided t-test on effect of trial type on maximum TE coefficient in a linear mixed effects model, followed by a Bonferroni correction for three comparisons. Shaded regions and error bars in all plots are 95% confidence intervals. Thin lines in violin plots and cue aligned averages represent averages for individual mice. For box plots in a,c,d,f,h, the centre point is the median, the lower and upper bounds are the first and third quartiles, and the whiskers extend from the box bounds to the most extreme value no further than 1.5 x interquartile range from the bounds.

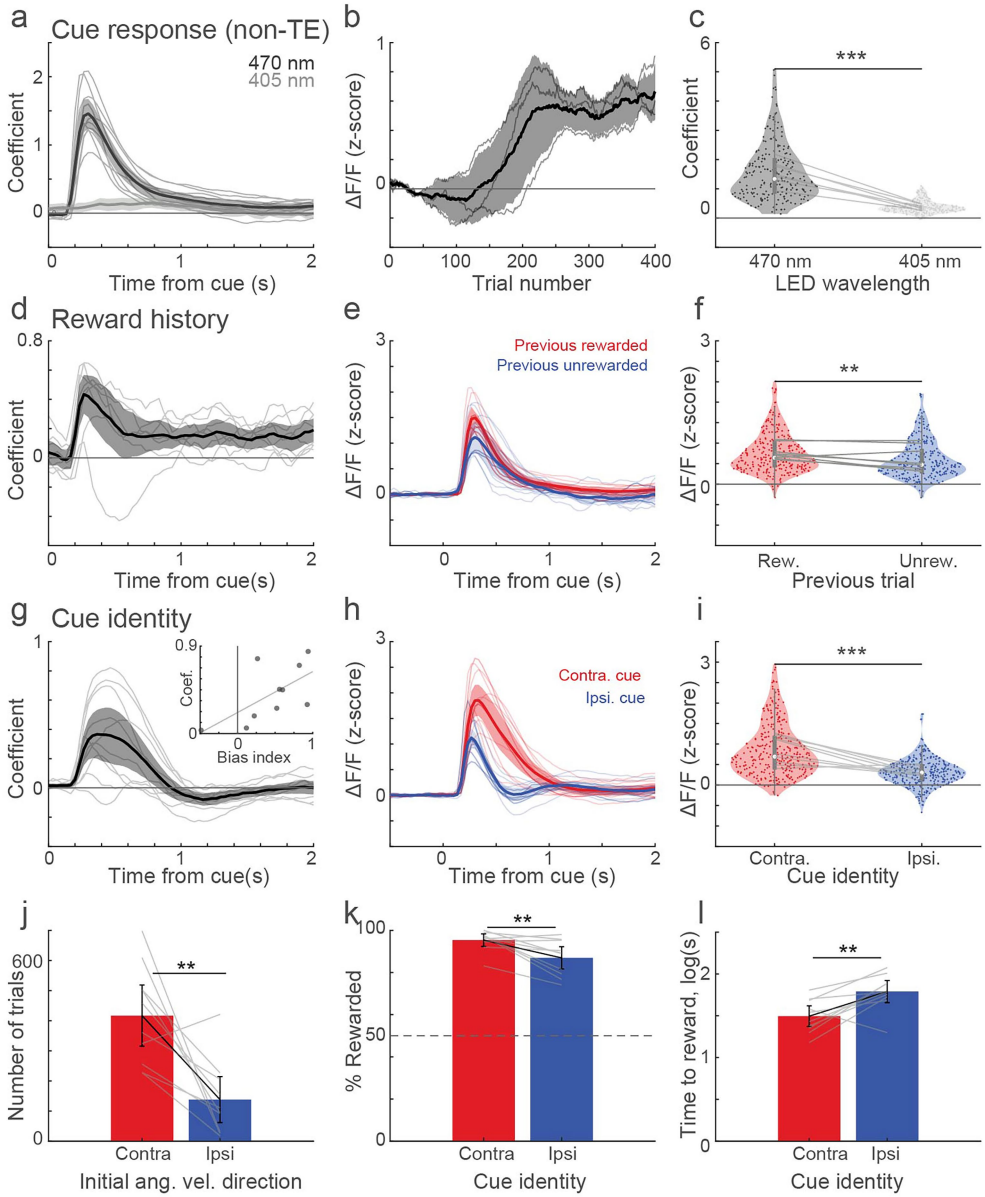


Extended Data Fig. 4 | Relationships of dopamine signaling with task-independent locomotion cannot account for trajectory error encoding. **a**, Average acceleration coefficients during running in directions contralateral (left) and ipsilateral (right) to the implant in a model describing ΔF/F as a function of continuous acceleration at different time lags during the intertrial interval (see Methods). Acceleration was calculated from the absolute value of the angular velocity for contralateral and ipsilateral running periods independently. Thin lines represent averages for each mouse, thick black line, average across mice (n = 193 sites, 6 mice) **b**, Left, average signed angular acceleration (+contralateral, − ipsilateral) aligned to cue onset on congruent (green) and incongruent (purple) contralateral cue trials, split into tertiles (dark to light colours indicates fast to slow) for one mouse. Accelerations on incongruent trials are oppositely signed for ipsilateral and contralateral trials, but TE coefficients are positive for both trial types (Extended Data Fig. 3g). **c**, Average absolute value of angular treadmill velocity aligned to cue onset across all trials (left) and trials where running was contralateral (middle) and ipsilateral (right) to the implant at cue onset during a non-directional, Pavlovian delay conditioning task (top left inset, see Methods). Trial averages are split into tertiles within 0.3 s before cue onset (dark to light colours indicate fast to slow trials). **d**, Average cue onset aligned ΔF/F for a single recording site on the trial types indicated in c. **e**, Violin plots showing maximum and minimum angular velocity coefficients across all sites for the trial types in c. For the boxplot, the centre point is the median, the lower and upper bounds are the first and third quartiles, and the whiskers extend from the box bounds to the most extreme value no further than 1.5 x interquartile range from the bounds. Each data point represents a single site. Inset text, proportion of fibers significant (two-sided t-test on model coefficients, $p < 0.05$ for 3 timepoints in a row within 1 s cue window, Bonferroni corrected). Shaded regions in all plots, 95% confidence intervals.

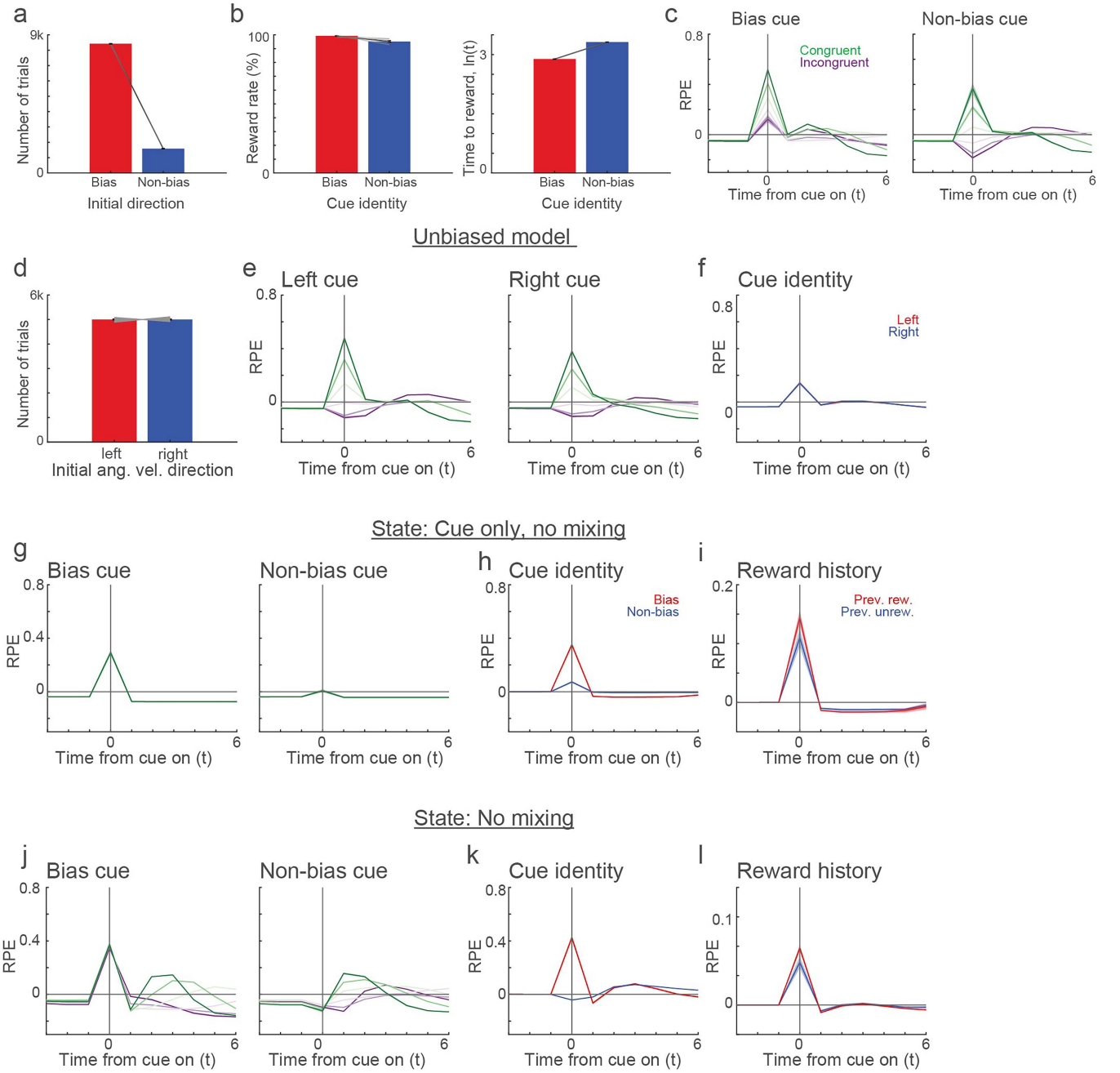


Extended Data Fig. 5 | See next page for caption.

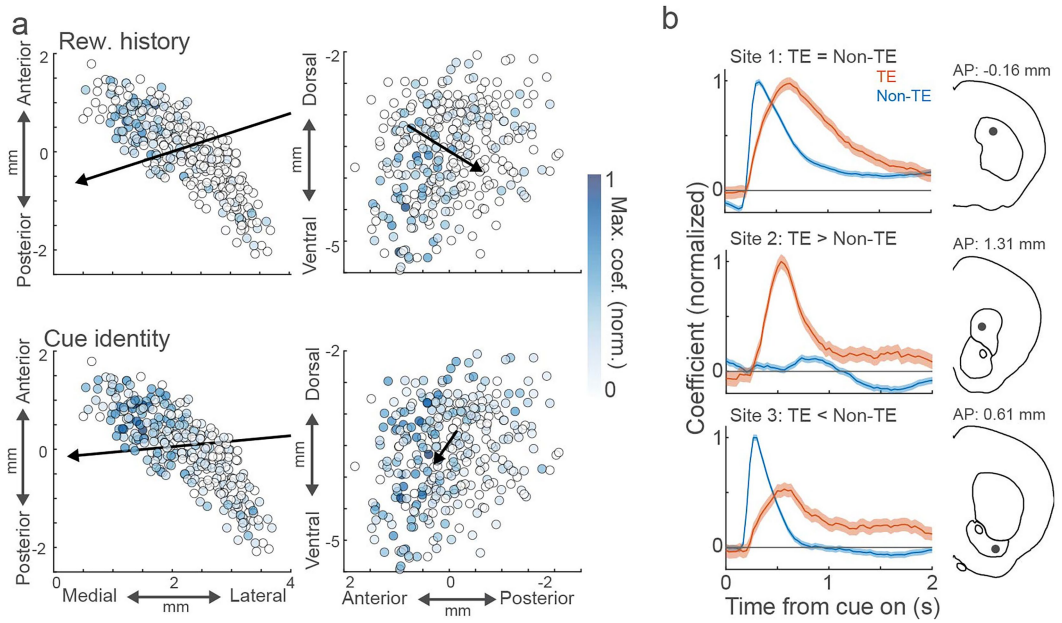
Extended Data Fig. 5 | Trajectory error encoding can be based independently on visual flow and locomotion. a, Schematic of the visual only task (vis-only) in which the associations between cue identity and rewarded cue position were the same as in the instrumental (loc + vis task, Fig. 1), but the mouse angular velocity was decoupled from movement of the visual cue. Cue direction and velocity on each trial were yoked to congruent and incongruent trials performed by separate mice (see Methods). **b**, Average lick indices on rewarded (blue) and unrewarded (red) trials for each training day for one example mouse showing successful discrimination of cue-direction associations after learning (n = 88, 92, 91, 69, 89, 91, 92, 89, 95 rewarded trials on days 1-9, respectively; n = 38, 39, 35, 26, 39, 39, 40, 38, 38 unrewarded trials on days 1-9, respectively.) **c**, Average lick indices across trials and days for rewarded (blue) and unrewarded (red) trials across mice. $t_3 = 72.50$, $p = 5.78 \times 10^{-6}$, n = 4 mice, two-sided paired t-test on the effect of trial type on lick index. **d**, Top, average congruence-signed cue angular velocity (+congruent, −incongruent) aligned to cue onset on congruent (green) and incongruent (purple) trials, split into thirds by cue velocity magnitude for one example mouse. Bottom, same as top but with congruence signed relative to the angular mouse locomotion velocity at cue onset based on cue-direction contingencies in the loc + vis instrumental task. **e**, Same as d but for mouse angular velocity. Note the decoupling of cue and mouse velocity. **f**, Average cue aligned ΔF/F for an example recording site on congruent (green) and incongruent (purple) trials based on the cue (top) or mouse (bottom) congruence-signed angular velocity, split into thirds by magnitude, as in d and e respectively. **g**, Top, Trajectory error coefficients aligned to cue onset based on the congruence-signed angular cue (blue) and locomotion (green) velocity averaged across sites and mice with a linear mixed effects model. Bottom, Violin plot of the maximum visual and locomotion based TE coefficients (t-stats) for all fibers and mice $t_{230} = -8.93$, $p = 1.4 \times 10^{-16}$, n = 116 sites, 4 mice, two-sided t-test on main effect of TE type in a linear mixed effects model. **h**, Top, trajectory error coefficients aligned to cue onset based on congruence-signed angular velocity of the cue in the vis-only task (blue) and the loc + vis task (orange) averaged across sites and mice with a linear mixed effects model. Bottom, violin plot of the maximum TE coefficients (t-stats) for all sites and mice. $t_{230} = -0.001$, $p = 0.999$, two-sided t-test on main effect of task type in a linear mixed effects model, n = 116 sites, 4 mice. **i**, Schematic of the locomotion only task (loc-only) in which task rules were the same as loc+vis but the visual cue remained stationary at the centre of the screens. **j**, Average reward rates in the loc-only (pink) and loc + vis tasks (orange), for sessions above criterion. Mouse 1: $X^2 = 1.36$, $p = 0.49$, n = 762 and 535 trials for loc+vis and loc-only, respectively; mouse 2: $X^2 = 2.64$, $p = 0.21$, n = 792 and 786 trials for loc+vis and loc-only, respectively; Chi-squared test for comparing two proportions, p-values bonferroni corrected for 2 mice. **k**, Histogram of times to reward for each task. **l**, Average congruence-signed mouse angular velocity (left) and ΔF/F for a representative fiber (right) aligned to cue onset for congruent (green) and incongruent (purple) trials, split into thirds by magnitude, during the loc-only task. **m**, Trajectory error coefficients for the loc-only (pink) and loc + vis tasks (orange) averaged across fibers and mice (n = 79 sites, 2 mice). **n**, Violin plot of the max loc-only (pink) and loc + vis (orange) TE coefficients for all fibers and mice. $t_{156} = 1.70$, $p = 0.093$, two-sided t-test on main effect of task type in a linear mixed-effects model, n = 79 sites, 2 mice. Shaded regions and error bars in all plots are 95% confidence intervals. Thin lines in all plots are averages for individual mice. For the boxplots in g,h,n, the centre point is the median, the lower and upper bounds are the first and third quartiles, and the whiskers extend from the box bounds to the most extreme value no further than 1.5 x interquartile range from the bounds.



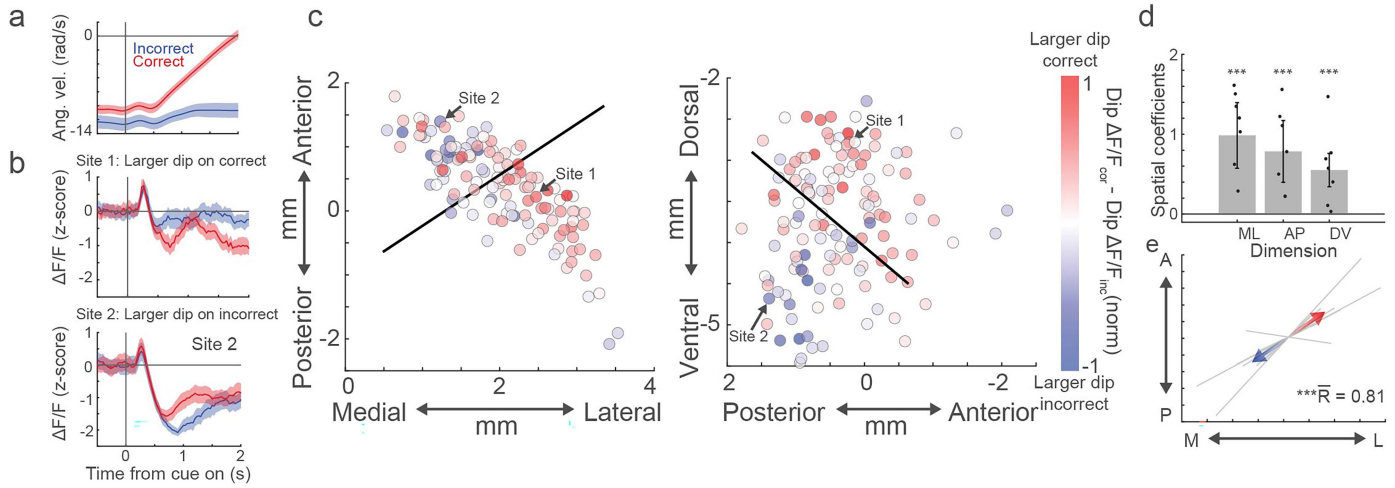
Extended Data Fig. 6 | Cue-evoked dopamine release encodes cue-value RPEs. a, Non-TE DA cue responses averaged across sites and mice using a linear mixed-effects model (470: $df = 315$, $n = 316$ sites, 10 mice; 405: $df = 230$, $n = 231$ sites, 7 mice). **b**, Maximum cue-evoked z-scored ΔF/F during early learning for a subset of mice ($df = 79$, $n = 80$ sites, 3 mice). **c**, Violin plot of maximum non-TE coefficients for mice with quasi-simultaneous 470 nm and 405 nm recordings (two-sided t-test on wavelength effect: $p = 1.33 \times 10^{-18}$, $t_{460} = -9.19$, $n = 231$ sites, 7 mice). **d**, DA reward-history coefficients averaged as in **a** for mice with >30 unrewarded trials (470: df = 260, $n = 261$ sites, 8 mice; 405: $df = 175$, $n = 176$ sites, 5 mice). **e**, Z-scored ΔF/F aligned to cue onset following rewarded (red) or unrewarded (blue) trials, averaged across sites and mice ($df = 260$, $n = 261$ sites, 8 mice). **f**, Violin plot of maximum ΔF/F within 1 s from cue onset for trials in e (two-sided t-test on reward-history effect: $p = 0.0052$, $t_{520} = -2.80$, $n = 261$ sites). **g**, DA cue-identity coefficients averaged as in **a** (mixed-effects intercept: $df = 315$, $n = 316$ sites, 10 mice). Inset, maximum cue-identity coefficients vs contralateral bias index (Pearson $r = 0.63$, $p = 0.025$, n = 10 mice). **h**, Z-scored ΔF/F for cues signaling contralateral (red; bias cue in 9/10 mice) or ipsilateral (blue) directions, averaged as in **e**. **i**, Violin plot of maximum cue-identity coefficients (two-sided t-test on cue-identity effect: $p = 6.77 \times 10^{-7}$, $t_{630} = -5.02$, $n = 316$ sites, 10 mice). **j**, Mean number of trials with initial velocity directed contralaterally (red) or ipsilaterally (blue) to implant (two-sided paired t-test: $p = 0.0037$, $t_9 = -3.89$, $n = 10$ mice). **k**, Reward rate for contralateral (red) vs ipsilateral (blue) cue trials (two-sided paired t-test: $p = 0.0087$, $t_9 = 3.33$). **l**, Time to reward for contralateral (red) vs ipsilateral (blue) cue trials (two-sided paired t-test: $p = 0.0066$, $t_9 = -3.51$). Shaded regions and error bars denote 95% confidence intervals; thin lines, individual-mouse averages. Box plots: centre, median; box bounds, first and third quartiles; whiskers, extremes value less than 1.5 x interquartile range from bounds.



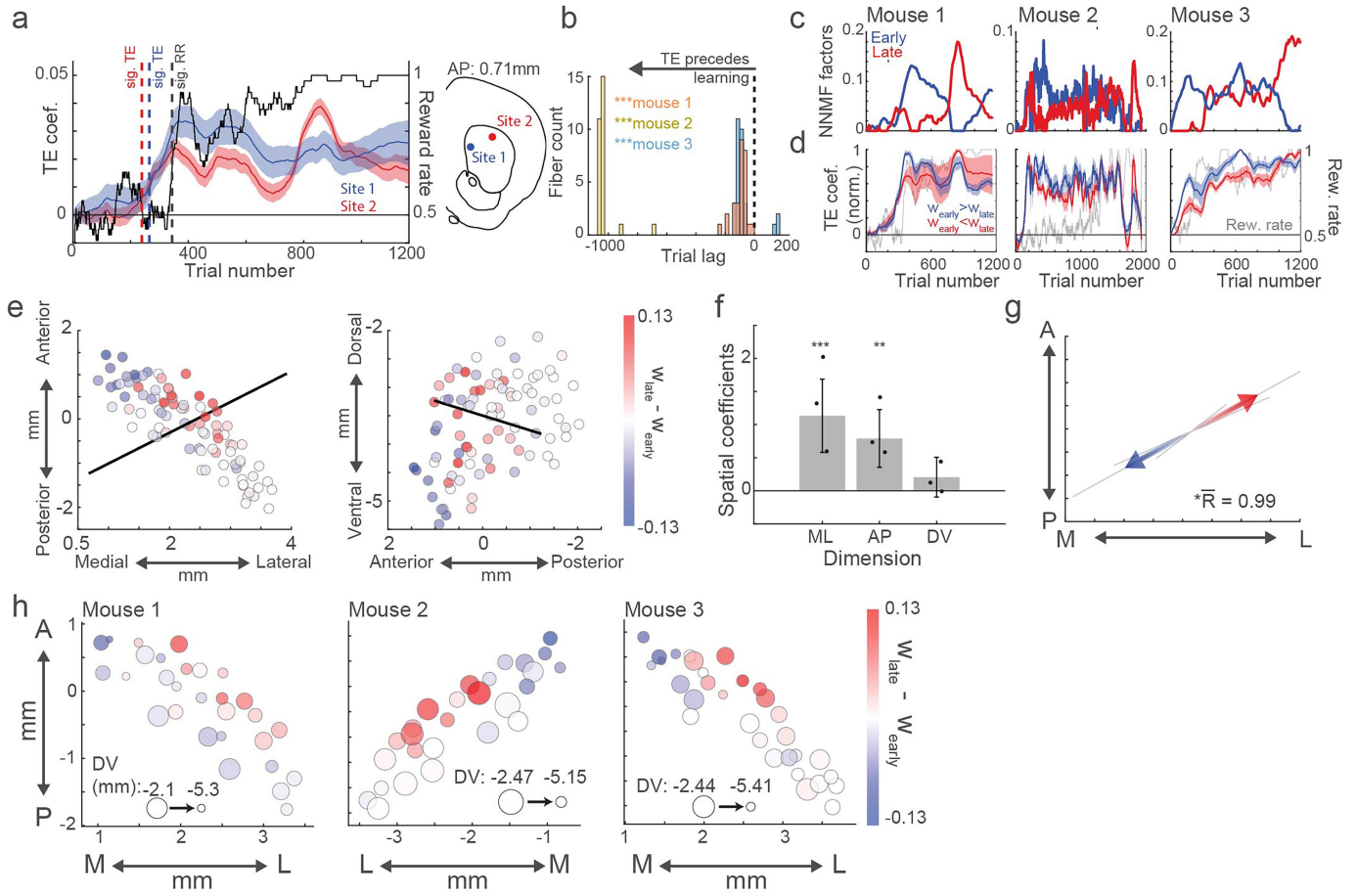
Extended Data Fig. 7 | RPE encoding of TEs requires mixed state-space representations of sensorimotor variables. a, Initial running direction bias introduced into the tdRPE model to mimic mouse behaviour (Fig. 3j). Bars represent average trial numbers for bias and unbiased directions across 10 model simulations. **b**, Average reward rates (left) and times to reward (right) for cues associated with the pre-cue bias and non-bias initial running directions (compare to mouse behaviour data in Fig. 3k–l). **c**, Average simulated RPEs aligned to cue onset ($t = 0$) for trials presented with the cue associated with the bias (left) or non-bias (right) directions, for congruent (green) and incongruent (purple) trials split into thirds by simulated initial angular velocity magnitude. **d**, Average trial numbers for each cue for a model with no initial direction bias, presented as in **a**. **e**, Same as c for an un-biased model, split by cues associated with left and right directions. **f**, Average simulated RPEs aligned to cue onset on trials presented with the cue associated with the right (red) or left (blue) direction in an un-biased model. **g-i**, Results from a tdRL model (with the directional bias in a) in which the state only includes the cue identity variable with no mixing. **g**, Same as c. **h**, Same as f. **i**, Same as f with trials split by whether previous trial was rewarded (red) or unrewarded (blue). RPEs are larger for the cue associated with the initial running bias and following previously rewarded trials, consistent with the mouse DA data and cue value RPE encoding (Fig. 3d–i). **j-l**, Results from a model (with the directional bias in a) in which the state includes all variables (Fig. 3b) but with no mixing. RPEs are plotted as in **g-i**. Shaded regions and error bars in all plots are 95% confidence intervals.



Extended Data Fig. 8 | Spatial maps and model coefficients for cue value measures across sites. a, Horizontal (left) and sagittal (right) maps of the maximum reward history (top) and cue identity (bottom) coefficients for each recording site (bubble) across mice, normalized to the maximal coefficients across all recording sites. Empty bubbles, no significant coefficients. Black arrow is the axis of maximal variation, pointing in the direction of increasing magnitude based on the spatial coefficients in Fig. 5h. **b**, Trajectory error (orange) and non trajectory error (cue value, Fig. 5; blue) coefficients for three example sites in the same mouse (locations at right). Shaded regions are 95% confidence intervals around the FIR coefficient estimates (centre line). Brain schematics were adapted from the Allen Mouse Brain Common Coordinate Framework (CCFv3) (https://atlas.brain-map.org/).



Extended Data Fig. 9 | Spatial organization of trial-trial performance encoding in trajectory error signaling. a, Angular velocity, signed by congruence, aligned to cue onset averaged across correct (blue) and incorrect (red) incongruent trials for one example mouse. **b**, ΔF/F aligned to cue onset averaged across correct (blue) and incorrect (red) trials for two example recording sites in the same mouse. Recording site locations indicated in c. **c**, Horizontal (top) and sagittal (bottom) maps of the difference in dip magnitude between correct and incorrect trials. Colour of each bubble (recording site) represents the difference, with positive values indicating a larger dip for correct trials. **d**, Spatial coefficients representing the magnitude and direction of the variation in the difference in cue-evoked dip magnitude between correct and incorrect incongruent trials at recording sites with a significant dip on either trial type (ML: $t_{127} = 4.74$, $p = 1.66 \times 10^{-5}$; AP: $t_{127} = 4.01$, $p = 3.06 \times 10^{-4}$; DV: $t_{127} = 5.21$, $p = 2.25 \times 10^{-6}$; $n = 131$ sites, 10 mice, two-sided t-test on effect of location on dip magnitude difference in a LME, Bonferroni corrected). **e**, Vector showing the axis of maximal variation in dip magnitude difference, based on the coefficients in d. $\underline{R}$, Rayleigh's coefficient, $S^* = 22.38$, $p = 5.43 \times 10^{-5}$. Thin lines represent individual mice. Shaded regions and error bars in all plots are 95% confidence intervals around the mean (**a,b**) or LME coefficient estimates (**d**).



Extended Data Fig. 10 | Spatial and temporal variations in TE expression across learning. a, Dynamic regression model estimates of trial-by-trial reward rates (black) and TE coefficients for two example fibers across all trials and sessions through learning (see Methods). Dashed lines indicate the trials in which the TE coefficients and reward rates reached significance ($p < 0.01$, shuffle test). Fiber locations indicated in the coronal atlas schematic. Black line, reward rate, averaged with a moving average window of 50 trials. **b**, Histogram of lags between the earliest trial where TE and reward rate reached statistical significance for each fiber in the three included mice. Colours indicate different mice. Negative values indicate that TE becomes significant prior to reward rate. Mouse 1: $t_{27} = -10.9$, $p = 7.0 \times 10^{-11}$, $n = 28$ sites; Mouse 2: $t_{27} = -76.2$, $p = 1.4 \times 10^{-32}$, $n = 28$ sites, Mouse 3: $t_{29} = -5.4$, $p = 2.6 \times 10^{-5}$, $n = 30$ sites, t-test on difference in trials between significant trajectory error encoding and significant performance, p-values Bonferroni corrected for 3 mice. **c**, Non-negative matrix factors (NNMF) captured from trial-by-trial TE trajectories (estimated with dynamic regression) across sites within each mouse representing early (blue) and late (red) TE increases across training for each of the three mice recorded across learning (see Methods for details). **d**, TE coefficients across trials averaged across recording sites with a larger weight for the early (blue) or late (red) component and normalized to the maximum TE coefficient, for each mouse. Grey line, reward rate averaged with a moving average window of 50 trials. **e**, Horizontal (left) and sagittal (right) plane maps of the differences in the weights for the late and early components of the NNMF for each recording site (bubble) across mice (n = 88 fibers, 3 mice). **f**, Spatial coefficients representing the magnitude and direction of the variation in the difference between early and late NNMF weights in e. Data points represent individual mice. ML: $t_{84} = 4.1$, $p = 3.11 \times 10^{-4}$; AP: $t_{84} = 3.6$, $p = 0.0016$; DV: $t_{84} = 1.4$, $p = 0.54$, two-sided t-test on spatial coefficients, p-values bonferroni corrected for 3 dimensions. Error bar, 95% confidence interval around LME model coefficients. **g**, Vector showing the axis of maximal variation in weight difference, based on the coefficients in g. $\underline{R}$, Rayleigh's coefficient, $S^* = 9.92$, $p = 0.02$. Thin lines represent individual mice. **h**, Horizontal plane maps of the differences in the weights for the late and early components of the NNMF for each mouse as in e. The bubble sizes indicate the DV position (increasing size indicates more dorsal positions). Shaded regions in all plots are 95% confidence intervals. Brain schematics were adapted from the Allen Mouse Brain Common Coordinate Framework (CCFv3) (https://atlas.brain-map.org/).

nature portfolio

Corresponding author(s): Mark Howe

Last updated by author(s): 10/31/2025

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	Data collection was performed using custom MATLAB code, and ran using MATLAB 2020b and Hamamatsu software.
Data analysis	All data analysis was performed using custom MATLAB code Code used for analysis will be deposited on a publicly available Zenodo.

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

All data used to generate figures will be made publicly available through Zenodo. Links will be provided with the final manuscript proof.

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	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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	Our sample sizes are consistent with or larger than similar publications using comparable methods (see refs 25, 59)
Data exclusions	Recordings from mice that were unable to learn based on predefined criteria were excluded. Recordings from fibers that were located outside of the striatum or did not have sufficient SNR based on predefined criteria were excluded. More details are described in the methods section.
Replication	Findings were successfully replicated across recording sites and mice as verified using linear mixed effects models
Randomization	No comparisons across groups were performed. All analyses was performed on neural signals within groups.
Blinding	Blinding was not relevant to our study because no cross-group comparisons were performed that would necessitate this.

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

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	Male and female C57BL6 mice (14 total), ages 9-20 weeks at implantation were used across experiments
Wild animals	No wild animals were used in this study
Reporting on sex	All findings in the study apply to both sexes. Across all experiments, 8 male and 6 female mice were used. We have indicated in Table 2 the breakdown of groups for individual behavior tasks.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All animal use was approved and monitored by the Boston University IACUC committee

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

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A