Article

Neural basis of compositional control

https://doi.org/10.1038/s41586-026-10896-8

Received: 14 March 2025

Accepted: 8 July 2026

Published online: 12 August 2026

Check for updates

Assia Chericoni[1,15], Justin M. Fine[1,15], Taha S. Ismail[1], Gabriela Delgado Salazar[2], Melissa C. Franch[1], Elizabeth Mickiewicz[1], Ana G. Chavez[1], Eleonora Bartoli[1,3], Danika L. Paulo[1], Vaishnav Krishnan[3,4,5,6], Mohamed Hegazy[5], Alica M. Goldman[5], Lu Lin[5], Garrett P. Banks[1,7], Nisha Giridharan[1], Mohammed Hasen[8], Nicole R. Provenza[1,3,4,9,10], Andrew Watrous[1], Seng Bum Michael Yoo[11,12,13], Sameer A. Sheth[1,3,4,6,10,14,16] & Benjamin Y. Hayden[1,3,4,10,16] ✉

Naturalistic goal-directed behaviour often involves continuous actions directed at dynamically changing goals[1–3]. Just as microeconomics serves as a rigorous foundation for discrete choices, control theory can serve as a foundation for understanding choice in continuous ones[3,4]. In continuous contexts, behaviour is composed of blends of goals, and the closest analogue to choice is a strategic reweighting of goal-specific control policies[5,6]. Here, to understand the algorithmic and neural bases of continuous choice, we examined behaviour and brain activity in humans performing a continuous prey-pursuit task[7]. Using a newly developed control-theoretic decomposition of behaviour, we find that pursuit strategies are well described by a meta-controller dictating a mixture of lower-level controllers, each linked to specific pursuit goals. Neurons in the anterior cingulate cortex predict major changes in policy blends, whereas hippocampal neurons encode and update the latent policy state supporting early planning. Meanwhile, orbitofrontal cortex activity is consistent with an encoding of the current value structure of the task, rather than policy switching. Together these results are consistent with a tripartite functional division in which hippocampus serves as a state-estimating controller, anterior cingulate cortex serves as a meta-controller, and orbitofrontal cortex provides a value context signal.

The neuroscientific study of choice is primarily based on discrete choices with static options. Yet many of the real-world decisions are continuous and interactive[1–3]. Consider a predator pursuing a fleeing prey[7,8]. At each moment, the predator chooses a single direction to travel, while anticipating the response of the prey and incorporating ongoing feedback (control policies). When there are multiple prey, the predator must also decide whether and when to adjust which goal, or blend of multiple goals, is pursued[5,9,10]. Changes in the composition of these goal blends are the closest analogue to choices in continuous contexts.

Optimal policy blending requires a meta-controller that can solve a non-trivial and nonlinear goal optimization problem[11,12]. In general, this problem is intractable, because a new nonlinear controller would have to be learned for each new set of goals[13]. However, one solution is compositional control, which builds on control theory, artificial intelligence and robotics[14]. In compositional control, nonlinear control is approximated via composition of individual linear controllers[14,15]. Each controller resolves the best actions for a single goal. Crucially, these controllers depend on internal state estimation to map and track the hidden dynamics of the environment. The behaviour of the system then reflects a blended weighting of multiple controllers, with a meta-controller monitoring shifts in goal priorities and adjusting the relative weighting of individual controllers accordingly[16,17]. Understanding whether and how the brain might implement compositional control can explain the neural mechanisms enabling humans to arbitrate between competing goals in dynamic environments and help us generalize choice to continuous contexts.

Brain regions coding for a policy weighting should have preferential access to information needed for creating mappings of goals (for example, goal distance) and goal-specific control policies[18,19]. This requirement aligns with modern theories of hippocampus (HPC) function[20–24]. Nonetheless, most studies view HPC as a static map-like resource for other regions that perform control-related computations, leaving its potential role in control underexplored.

Meanwhile, value comparison and control-demand computations are more consistently attributed to frontal circuits[21,25,26]. This function resembles meta-control, which involves integrating information about task context, goal priorities and expected outcomes to regulate control strategies[27–29]. Previous work implicates the anterior cingulate cortex (ACC) in precisely these meta-control functions[30–36]. Building

[1]Department of Neurosurgery, Baylor College of Medicine, Houston, TX, USA. [2]Department of Neuroscience, University of Minnesota, Minneapolis, MN, USA. [3]Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA. [4]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA. [5]Department of Neurology, Baylor College of Medicine, Houston, TX, USA. [6]Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, USA. [7]Department of Neurosurgery, Michael E. DeBakey VA Medical Center, Houston, TX, USA. [8]Department of Neurosurgery, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia. [9]Department of Bioengineering, Rice University, Houston, TX, USA. [10]Neuroengineering Initiative, Rice University, Houston, TX, USA. [11]Center for Neuroscience Imaging Research, Sungkyunkwan University, Suwon, South Korea. [12]Department of Biomedical Engineering, Sungkyunkwan University, Suwon, South Korea. [13]Department of Intelligence Health Precision Convergence, Sungkyunkwan University, Suwon, South Korea. [14]Gordon and Mary Cain Pediatric Neurology Research Foundation Laboratories, Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, USA. [15]These authors contributed equally: Assia Chericoni, Justin M. Fine. [16]These authors jointly supervised this work: Sameer A. Sheth, Benjamin Y. Hayden. ✉e-mail: benjamin.hayden@bcm.edu

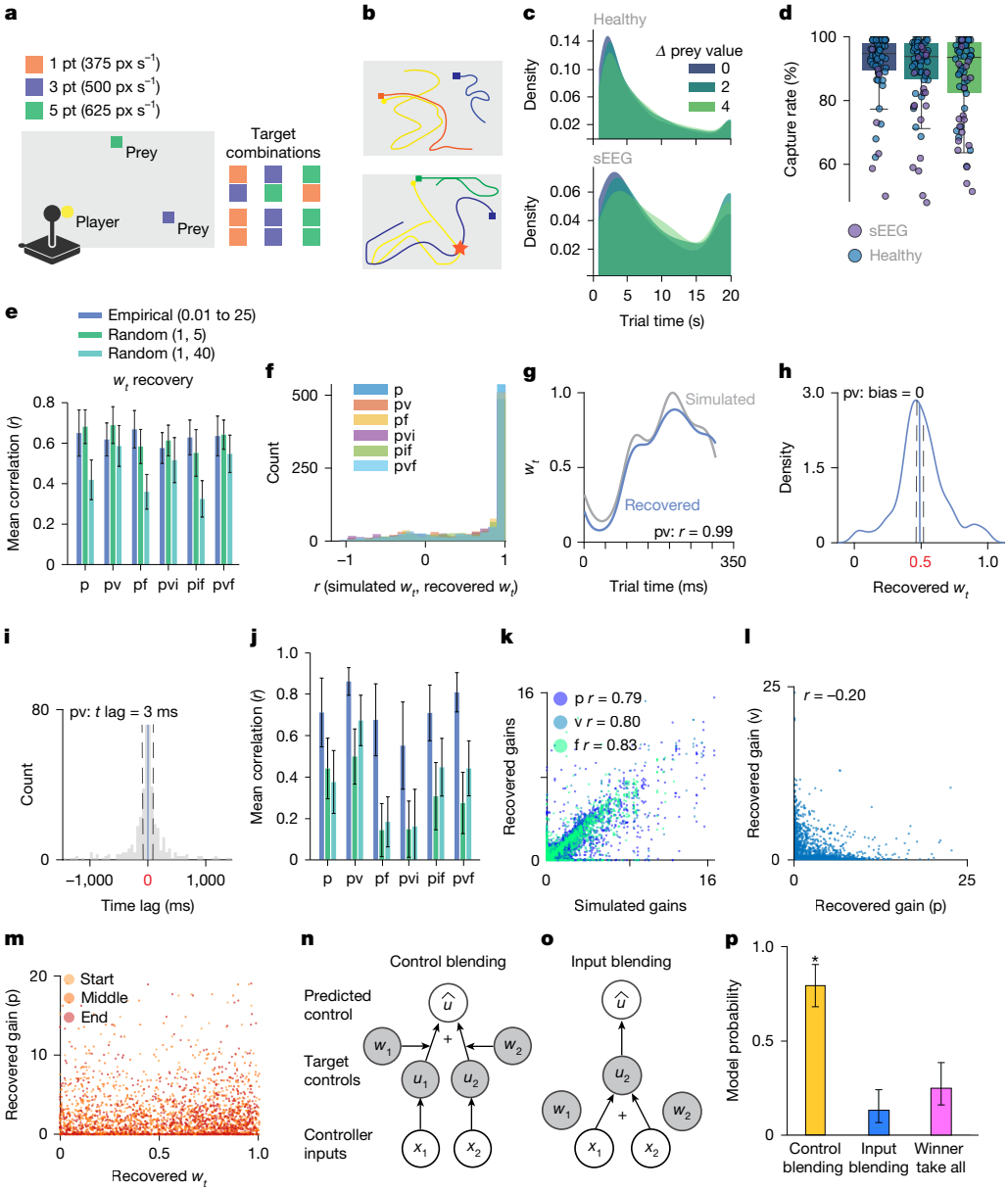


Fig. 1 | Continuous prey-pursuit task and control model validation. **a**, Schematic of the prey-pursuit task. Prey were displayed as squares and participants' avatar as a circle, with velocity determined by joystick deflection. Right, reward combinations across trial types. **b**, Example trajectories from two trials. **c**, Capture times across trial types. Δ denotes absolute difference between prey values. Results are shown for healthy ($\Delta = 0, 2$ or 4; $n = 50$) and sEEG ($\Delta = 0$, $n = 7$; $\Delta = 2$ or 4, $n = 19$) participants. **d**, Capture rate across trial types. Dots represent individual participants (sample sizes as in **c**). Box plots show the median (centre line), interquartile range (box) and 1.5× interquartile range (whiskers). **e**, Recovery of w_t across controller types and simulations regimes. Controller types: p (position), pv (position-velocity), pf (position-future position), pvi (position-velocity-integral position), pif (position-integral position-future position) and pvf (position-velocity-future position). Bars show mean Pearson correlation between simulated and recovered w_t; error bars indicate s.d. ($n = 30$ simulations per regime; 30 trials each). **f**, Distribution of w_t recovery correlations in the empirical simulation regime. **g**, Example w_t trajectory from a simulated trial. **h**, Distribution of the recovered w_t around the simulated switch point ($w_t = 0.5$); dashed lines indicate 95% confidence interval. **i**, Time-lag histogram between simulated and recovered w_t switch times; dashed lines indicate 95% confidence interval. **j**, Recovery of gain parameters across simulation regimes and controller types. Bars show mean Pearson correlation between simulated and recovered gains; error bars indicate s.d. (sample size as in **e**). **k**, Correlation between simulated and recovered controller gains (empirical simulation regime). **l**, Lack of correlation among recovered gains. **m**, Lack of correlation between recovered gains and w_t across trial epochs (all $r < 0.1$). **n**, Schematic of the control blending model. Shaded circles represent latent variables (w, u_1 and u_2); transparent circles represent observed variables (x_1, x_2 and $\hat{u}$). **o**, Schematic of the input blending model. **p**, Bayesian model comparison of candidate controller architectures ($n = 13$). Bars show mean posterior probability across participants; error bars indicate s.e.m. The asterisk denotes the model with highest posterior probability.

on this literature, we predicted that ACC neurons would encode policy switching in a value-dependent manner.

A meta-control system that reallocates control across goals also requires access to a representation of the current value structure of the task. The orbitofrontal cortex (OFC), which encodes the relative desirability of available options and the relationships between states and outcomes that define the value landscape of a task, can provide this information[26,37,38]. Beyond scalar value signals, OFC may organize this information into a cognitive map of task space, conceptually analogous to HPC cognitive maps, but emphasizing value and outcome

contingencies[39]. Consistent with this view, OFC representations may share structural similarities with HPC state representations, while differing in the information that they emphasize[40]. Thus, because OFC provides the value representations on which meta-control mechanisms operate, we predicted that its activity would primarily reflect the immediate value structure of the task.

We recorded brain activity in HPC, ACC and OFC in 19 patients undergoing intracranial monitoring for epilepsy (the stereoelectroencephalography (sEEG) cohort) and recorded behaviour only in a separate cohort of 50 healthy control individuals (the control cohort) performing a prey-pursuit task[7,41] (Fig. 1a–d, Extended Data Tables 1 and 2 and Methods). On each trial, the participant used a joystick to move an avatar freely in a two-dimensional area to pursue fleeing prey (Supplementary Video 1). Prey colours indicated their speed and point value (Fig. 1a,b). Three prey–value combinations were presented, ranging from equal-value prey to pairs with large value differences (Fig. 1a). Each trial ended with prey capture or after a maximum of 20 s. In the sEEG cohort, participants completed 81 ± 9 trials (mean ± s.d.); participants in the control cohort completed 227 ± 8 trials.

The average trial lasted 6.7 ± 5.8 s. Trials took longer when the difference in prey value increased (two-sided t-test across value conditions; all $P < 0.05$) owing to players' preference for larger valued faster prey (Fig. 1c). Capture rates were high across all conditions (87.6–90.8%; Fig. 1d).

Inferring continuous choice

In discrete decision-making tasks, identifying a participant's choices is trivial. In continuous decisions, inferring a participant's moment-to-moment decision strategy is more challenging[3]. To identify each participant's moment-to-moment control strategies, we must estimate both their constituent controllers and the mixture parameter that blends them (which we call w_t; Methods, 'Overview of control choice decomposition' equations). With 2 targets, w_t varies between 0 (by convention, we use this for the lower-valued target) to 1 (pursuing the higher-valued target); values in between reflect a blend of strategies.

Whereas control strategies can be assumed to correspond to pursuit of each prey item, the details of the control strategy cannot be assumed. Participants may minimize positional error, velocity error, integral of position error, estimated future position error or any combination of these. Thus, for example, estimating w_t under the assumption of a position-error controller would yield incorrect weights whenever the true controller depends on other variables. We therefore developed a model-based decomposition that leverages optimal control and compositionality theory to estimate controllers and their mixtures[17].

For each trial, we estimated w_t by modelling the participant's joystick control using a second-order linear dynamic system (that is, acceleration; Methods). This process involved optimizing separate controller gains for each target simultaneous with optimizing w_t. We repeated this process for all types and combinations of controller classes, ranging from the usage of current position errors between cursor and target, to integral error, or one step ahead future position error (see Methods and Extended Data Table 3 for a list of all classes). We used a Bayesian model-averaging framework to recover a per-trial estimate of w_t, computing a posterior probability weighted average of the w_t estimates obtained from all controller class combinations.

Validation of the blended control model

To validate our modelling approach, we performed recovery of ground-truth simulated w_t signals and controller gains. To ensure that recovery performance was evaluated both within and beyond the parameter range relevant to our data[42], we repeated the recovery analysis across three gain regimes: (1) random draws between 1 and 5; (2) random draws between 1 and 40; and (3) an empirical range matching parameters obtained from our dataset (0.01–25; units depend on controller class and are omitted for simplicity). Parameter recovery was quantified as the Pearson correlation between simulated and recovered parameters (30 trials over 30 simulations). We report results from the empirical regime here; the random-gain regimes yielded consistent results (see Extended Data Figs. 1–4, 6).

Our recovery analyses proceeded in three stages: recovery of the blending parameter w_t, recovery of the controller gains, and analysis of their joint interaction. Recovery of w_t was strong on average (mean $r = 0.57 \pm 0.10$; Fig. 1e), with correlations strongly skewed toward 1 (Fig. 1f,g). However, because high correlations can hide systematic biases in recovery[42] and our subsequent analyses build on switch points ($w_t = 0.5$), we quantified bias and timing offsets in recovered w_t. Across controller classes, recovery was unbiased (mean $w_t = 0.49$, 95% confidence interval [0.47,0.51]; Extended Data Fig. 1), and timing offsets were less than a single display frame (that is, 16.67 ms; median lag = 3.0 ms, 95% confidence interval [−96.4, 102.4] ms; Extended Data Fig. 1). Since our analyses focused on the controller class that best fit participants' behaviour, we examined recovery performance for that model: a position-velocity (pv) controller that minimizes both the distance to the prey and the mismatch in velocity between the participants and prey. This model was unbiased (mean $w_t = 0.5$, 95% confidence interval [0.49, 0.52]; Fig. 1h) and temporal lags were close to zero (median lag = 3.0 ms, 95% confidence interval [−80.5, 87.0] ms; Fig. 1i). Recovery of the controller gains was highest for the pv controller class ($r = 0.67 \pm 0.11$; Fig. 1j and Extended Data Fig. 2). For the remaining classes, the average recovery correlation was lower ($r = 0.42 \pm 0.15$; Fig. 1j,k and Extended Data Fig. 2).

We next assessed whether jointly estimating controller gains and w_t introduced identifiability issues[42]. In principle, smaller w_t values could compensate for larger gains and vice versa. To test this possibility, we computed their correlations at the beginning, middle and end of each simulated trial. These correlations were consistently low and stable across trial epochs (mean $r = 0.008 \pm 0.01$; Fig. 1m and Extended Data Fig. 3). In summary, compensatory trade-offs between w_t and gain parameters were absent and both sets of parameters were robustly identifiable.

We also found that w_t is largely unrelated to participant speed ($r = 0.14 \pm 0.18$, $P = 0.38$), relative distance ($r = 0.11 \pm 0.06$, $P = 0.53$) or relative speed ($r = 0.09 \pm 0.02$, $P = 0.25$) of the two prey across participants. Consistent with good parameter identifiability, correlations among recovered gains were generally modest (Fig. 1l and Extended Data Fig. 4), although higher correlations emerged in more complex models (pvf: $r = 0.42 \pm 0.20$; Extended Data Fig. 4).

To assess identifiability across controller classes, we performed full confusion analyses (Extended Data Figs. 5 and 6). The pv model was recovered above chance when it was the generative model (43% correct; 95% confidence interval [0.40 0.47], $P < 0.0001$; Extended Data Fig. 6), but also frequently fit data generated by the pvi and pvf models. Thus, because controller classes are partially nested, identification of the pv model does not imply that it reflects the true generative process.

To complement this analysis, we compared the blended controller model against two alternative strategies: all-or-none target selection and input blending (Fig. 1n,o; proof in Supplementary Note 1) and Bayesian model comparison showed that participant behaviour was best explained by the blended controller strategy (posterior model probability 84 ± 8% across participants; Fig. 1p). Further, because these models are partially nested, model selection should be interpreted in relative rather than absolute terms.

Signatures of continuous pursuit

We used the recovered w_t trajectories to understand participants' pursuit behaviour (example in Fig. 2a,b). We set a threshold for the pursuit of $w_t > 0.85$ and $w_t < 0.15$ for behaviours that we classified as pursuit of the higher-value and lower-value targets, respectively. All participants

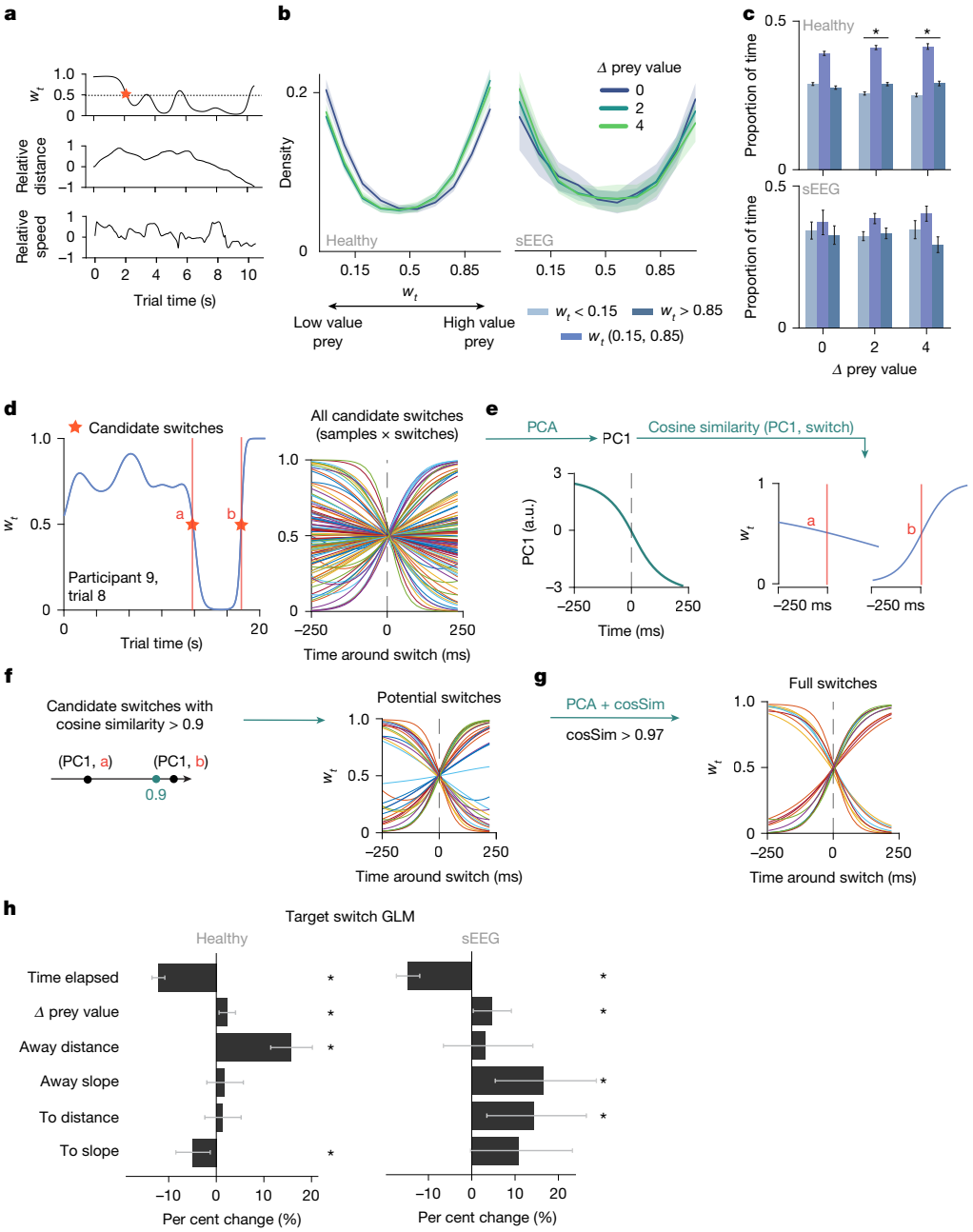


Fig. 2 | Target switch characterization. a, Example trial, showing the w_t trajectory (top), relative prey–participant distance (middle) and relative prey–participant speed (bottom). The red star marks a candidate target switch. **b**, Kernel density of participant-averaged w_t across reward conditions. Results are shown for healthy ($\Delta = 0$, 2 or 4; $n = 50$) and sEEG ($\Delta = 0$, $n = 7$; $\Delta = 2$ or 4, $n = 19$) participants. Shaded areas indicate s.e.m. across participants. **c**, Proportion of time spent pursuing each target or in a blended state across reward conditions. Data are mean ± s.e.m. Sample sizes are as in **b**. Asterisk denotes significance across conditions (two-sided paired t-test for healthy individuals: $\Delta = 2$, $P = 8.45 \times 10^{-4}$; $\Delta = 4$, $P = 3.68 \times 10^{-4}$). **d**, Candidate target switches identified as crossings of $w_t = 0.5$. **e**, Principal components analysis (PCA) of candidate switches; the first principal component (PC1) captures the characteristic sigmoidal switch trajectory. a.u., arbitrary units. **f**, Candidate switches were filtered by cosine similarity to the first principal component (threshold = 0.99). **g**, A second filtering step retained only complete switches. Sigmoids from high values to low values represent switches from the lower-value prey to the higher-value prey (and vice versa). **h**, Estimated per cent change in odds of switching prey from a Bayesian generalized linear model (GLM). Bars show posterior mean effects for healthy ($n = 50$) and sEEG ($n = 19$) participants; error bars indicate 95% credible intervals. Asterisks denote predictors with posterior probability greater than 0.95.

spent a substantial portion of time in a blended state ($0.15 < w_t < 0.85$): 36-41% of time in equal-value trials (Δ prey value = 0) and 37–43% in the different-value trials (Fig. 2b,c). As reward differences increased, time spent pursuing each option varied modestly (high 28–32%, low 26–33%). Overall, this pattern in behaviour demonstrates that pursuit decisions are not merely discrete but rather involve continuous reweighing of competing goals.

States underlying goal switching

We next identified target switches. We defined switches as periods in which w_t changed rapidly from less than 0.15 to greater than 0.85 or vice versa (Fig. 2d–g and Methods). Across both cohorts, this procedure yielded 13,357 candidate switches, of which 48% (6,366) were retained as switches for subsequent analysis (equal-value condition:

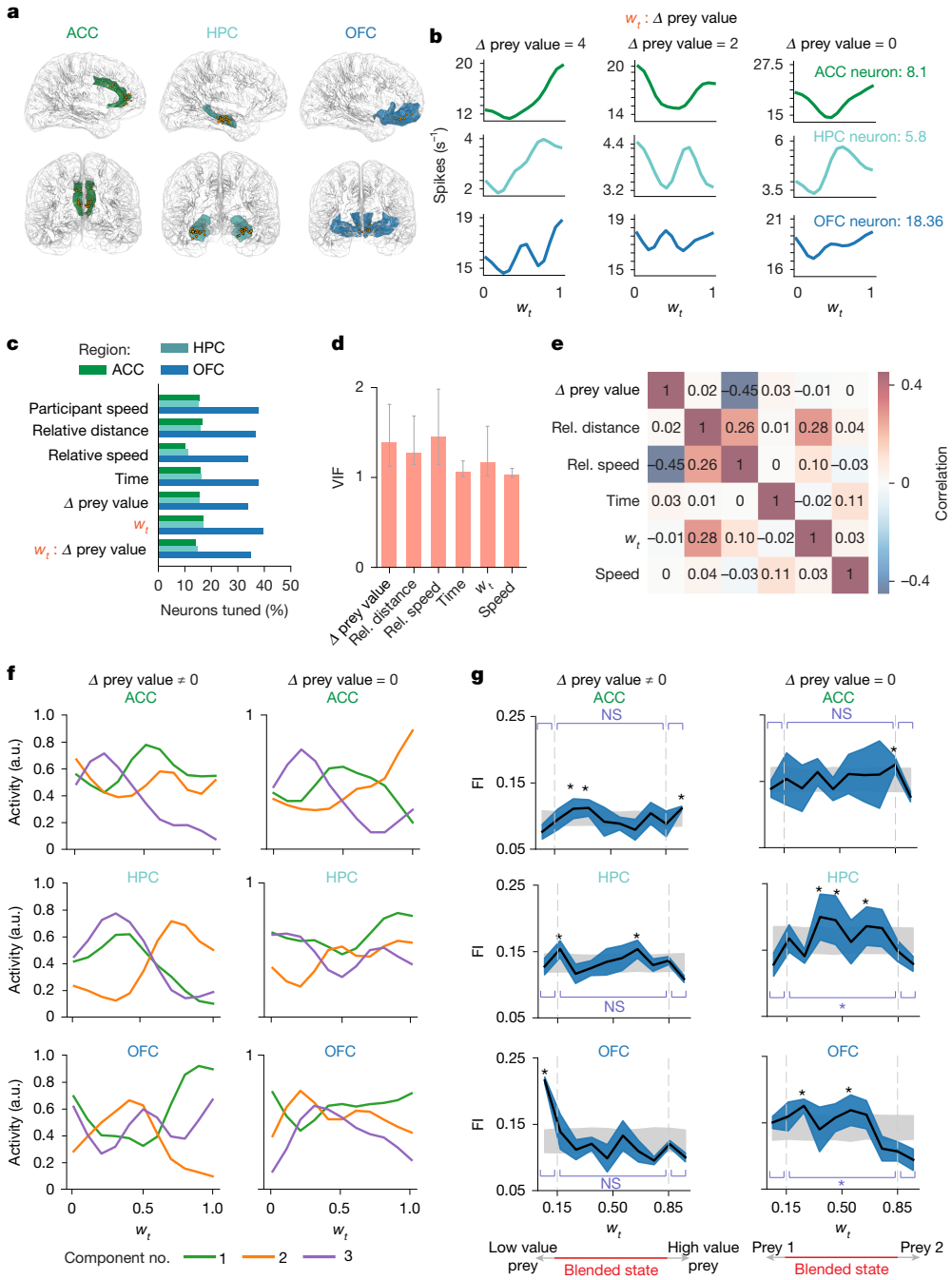


Fig. 3 | Neurons in the ACC, HPC and OFC are differentially tuned to w_t. **a**, Recording sites from the ACC (left), HPC (middle) and OFC (right). Dots represent electrodes across 19 patients. **b**, Exemplar tuning curves to w_t across reward conditions (columns) and brain regions (rows), estimated with the Poisson GAM. **c**, Percentage of neurons tuned to each variable across brain regions, as estimated with the GAM. **d**, Variance inflation factors (VIFs) assessing multicollinearity among GAM predictors. Bars indicate mean VIF across individuals ($n = 19$); error bars indicate the minimum-maximum range. All VIFs are lower than 2, indicating minimal multicollinearity and stable coefficient estimates. Rel., relative. **e**, Pairwise Pearson's correlations among predictors, showing low to moderate correlations. **f**, Non-negative matrix factorization components showing low-dimensional tuning bases to w_t for each brain region (rows) and reward condition (columns). **g**, Cross-validated FI with mean (black line) and s.e.m. (blue area). Grey shading represents the permuted distribution (95% confidence interval). Black asterisks denote bins with significant FI (one-sided permutation test, FDR-corrected across w_t bins; exact statistics are reported in Results). Purple brackets indicate the comparison between cumulative blended and committed states; purple asterisks denote significant differences (one-sided permutation test). The comparison varied significantly across reward conditions in HPC ($\Delta FI = 0.044$, $P = 0.001$) and OFC ($\Delta FI = 0.06$, $P = 0.001$), but not ACC. NS, not significant.

30.5%; Δ prey = 2: 46.3%; Δ prey = 4: 23.2%; Methods). Of these, half (52%) of switches were low to high; the remainder (48%) were high to low. Notably, switches did not occur simply when the future prey was closer than the currently pursued prey. At the switch point, the mean (z-scored) distance to the current and future prey did not differ (low-to-high switches: 0.76 versus 0.79; $P = 0.30$; two-sided t-test; high-to-low switches: 0.79 versus 0.76; $P = 0.29$).

Thus, to determine factors that lead to changes in propensity for target switching we used a hierarchical Bayesian logistic model (percentage change in odds of switching; Fig. 2h and Methods). In both cohorts,

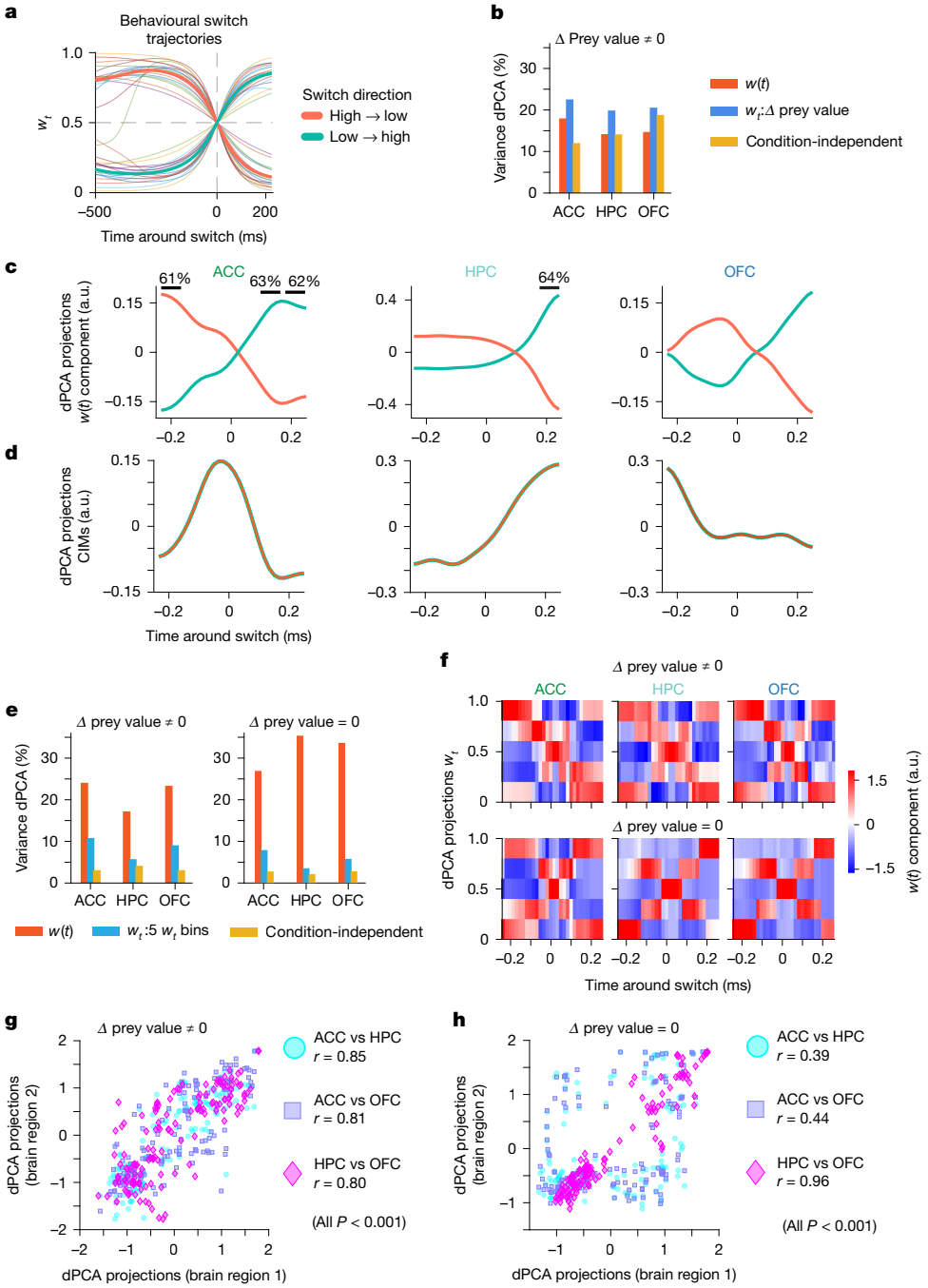


Fig. 4 | Low-dimensional controller blending dynamics during goal switch. **a**, Controller blending dynamics extracted during target switches and aligned to switch time (Fig. 2d–g and Methods). Orange traces denote switches from higher-value to lower-value prey, and green traces denote switches from lower-value to higher-value prey, averaged across participants. Thin lines represent the averaged trajectories per participant. **b**, Percentage of variance explained by each dPCA component. Here dPCA was configured to separate switch types. **c**, Neural population activity projected onto the leading w_t dPCA component. Black dashed lines indicate time bins with decoding accuracy significantly above chance; values above the dashed lines report decoding accuracy. Green and orange traces are as in **a**. **d**, Neural population activity projected onto the condition-independent dPCA component. Green and orange traces overlap. CIMs, condition-independent modes. **e**, Percentage of variance explained by each dPCA component. Here dPCA was configured to separate w_t representations across five bins for different prey value condition (left) and equal prey value condition (right). **f**, Neural population activity projected onto the leading w_t dPCA component for each brain region (columns) and reward condition (rows). Heat maps show mean dPCA projections across five w_t bins as a function of time relative to the switch (colour indicates component magnitude). **g**, Pairwise correlations between dPCA projections across brain regions in the different prey value condition. Each point represents a matched time w_t bin across regions; *r* denotes Pearson's correlation coefficient. Significance was assessed using two-sided Pearson correlation tests (ACC–HPC: $P = 7.81 \times 10^{-44}$; ACC–OFC: $P = 1.45 \times 10^{-35}$; HPC–OFC: $P = 4.76 \times 10^{-35}$). **h**, Same as **g**, but for the equal prey value condition (ACC–HPC: $P = 8.42 \times 10^{-7}$; ACC–OFC: $P = 2.39 \times 10^{-8}$; HPC–OFC: $P = 5.40 \times 10^{-86}$).

time spent pursuing a given target decreased the likelihood of switching (sEEG: −14.7%; control: −12.0%). Prey value difference increased likelihood of switching (sEEG: 4.7%; control: 2.1%). The effects of speed and distance (both from and to the prey) showed some variability across cohorts, as expected given behavioural heterogeneity (Fig. 2h). Switching behaviour was also influenced by interactions among predictors and

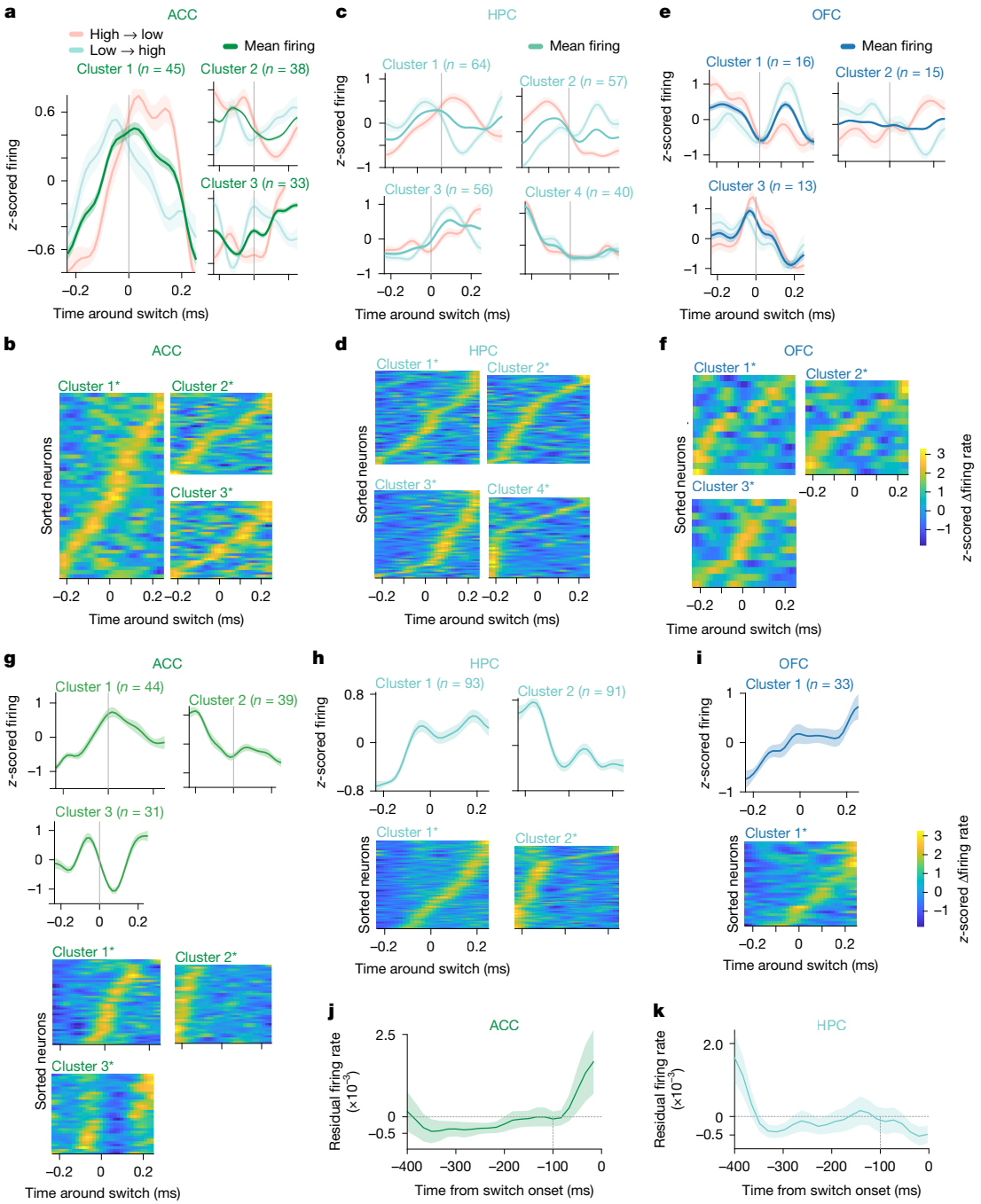


Fig. 5 | Neural ramping during changes of mind. a,**c**,**e**, Mean cluster firing rates in ACC (**a**), HPC (**c**) and OFC (**e**) aligned to switch onset in the different-reward condition. Lines show mean *z*-scored firing in each cluster; shaded areas indicate s.e.m. across neurons. **b**,**d**,**f**, *z*-Scored cluster selectivity for ACC (**b**), HPC (**d**) and OFC (**f**) in the different-reward condition, with neurons sorted by the switch-aligned time of maximal difference between switch and control activity. Asterisks denote clusters in which firing during switch windows significantly differed from control windows (two-sided permutation test, all $P < 0.01$). **g**–**i**, Same as panels **a**–**f**, but for the equal-reward condition. **j**,**k**, Mean residual firing rates in ACC (**j**) and HPC (**k**) aligned to switch onset, after regressing out behavioural covariates. Shaded areas indicate s.e.m. across neurons.

none of the model coefficients differed significantly between cohorts (Wald test, Bonferroni corrected, all $P > 0.05$). Together, these results show that target switches reflected a combination of multiple spatial, temporal and reward variables, rather than a single dominant factor.

Neural encoding of controller blending

We recorded single neurons while 19 patients performed the prey-pursuit task, including neurons from ACC (254 neurons, 15 participants), HPC (546 neurons, 18 participants) and OFC (103 neurons, 7 participants; Fig. 3a and Extended Data Table 2). These regions were recorded in at least five participants; no other region met this criterion. We used a Poisson generalized additive model (GAM; Methods) with a spline basis to estimate tuning for w_t, alongside task parameters and its interaction with prey value (Fig. 3b,c). We computed pairwise correlations and variance inflation factors among predictors (all less than 2), confirming that collinearity was minimal and did not affect the analysis (Fig. 3d,e). We found that neurons were tuned linearly to w_t in all the brain regions (16.9% in ACC; 17.2% in HPC and 39.8% in OFC; Fig. 3c). A substantial portion of neurons showed nonlinear mixing of w_t

Table 1 | Mixed-effects model results for the unequal-value condition (Δ prey value ≠ 0)

Δ prey value ≠ 0	Predictors significance (*P* values for full model)									
Brain region	**Time from switch onset (ms)**	**Time until switch**	**Time elapsed**	**Δ prey value**	**Away distance**	**To distance**	**Away slope**	**LR χ^2**	**LR (*P* value)**	**ΔAIC**
ACC	−400	<0.001	0.60	<0.001	<0.001	0.29	<0.001	15.18	<0.001	13.18
HPC		<0.001	0.002	0.07	0.002	<0.001	0.07	12.49	<0.001	10.50
ACC	−100	0.03	0.84	0.02	0.30	0.68	0.06	4.77	0.03	2.77
HPC		0.81	0.02	0.28	<0.001	<0.001	0.03	0.06	0.81	−1.94

Predictor *P* values from the full model are reported for two pre-switch windows (−400 ms and −100 ms) and reflect two-sided Wald *z*-tests on fixed-effect coefficients. Model comparisons between full and reduced models (lacking time until switch) were assessed using likelihood-ratio (LR) χ^2 tests and ΔAIC ($AIC_{reduced} - AIC_{full}$). Positive ΔAIC values indicate a better fit of the full model.

with relative prey value (14.2% in ACC; 14.8% in HPC and 35.0% in OFC, Fig. 3c). We also found neurons tuned to all auxiliary variables (Fig. 3c).

By examining how Fisher information (FI; Methods) is distributed across w_t, we can compare how sensitively neural populations from each brain region can discriminate changes in the policy value. For example, if a population is chiefly concerned with the value of a pursuit strategy, the FI will be skewed toward extreme w_t values (for example, pursuing higher-value prey). Conversely, a more similar FI around low and high w_t indicates a more balanced strategy. We first visualized the population tuning to w_t using a non-negative matrix factorization for dimensionality reduction (Fig. 3f). In ACC, we found a skew towards high-value prey, whereas HPC showed more balanced selectivity. OFC components differentiated low and high w_t when prey values differed, with weaker differentiation in the equal-value condition (Fig. 3f). Thus, population tuning was not uniformly distributed across the policy space but differed across brain regions.

The non-negative matrix factorization analysis provides visualization; we next quantified this by examining FI in separate permutation tests for each w_t bin. In ACC, we found significant w_t information for high prey values across reward conditions (one-sided permutation test, both false discovery rate (FDR)-corrected, Benjamini–Hochberg adjusted $P = 0.01$); whereas FI was high for blended states only when prey value differed (both FDR-corrected $P = 0.005$; Fig. 3g). In the HPC, FI was high for blended states (one-sided permutation test, all FDR-corrected $P = 0.02$; Fig. 3g). In OFC, FI peaked at low w_t values when prey values differed ($w_t < 0.15$, one-sided permutation test, FDR-corrected $P = 0.01$), whereas in the equal-value condition, it was distributed across blended states ($0.15 < w_t < 0.85$, both FDR-corrected $P = 0.02$; Fig. 3g). HPC and OFC had higher FI in blended states than in committed states in the equal-value condition (one-sided permutation test, both $P = 0.001$; purple brackets in Fig. 3g), but not in the unequal-value condition ($P = 0.51$ and $P = 0.99$ respectively). This difference across reward conditions was itself significant (HPC: $P = 0.001$; OFC: $P = 0.001$). By contrast, ACC showed no significant differences within and across conditions (all $P > 0.05$; Fig. 3g), suggesting that its FI is not dominated by a global difference between blended and commitment states.

Thus, neural information about w_t was differentially organized across brain regions. HPC and OFC preferentially represented blended policy when the alternatives were equally valuable, whereas the ACC maintained a value-biased w_t representation across value conditions.

Neural dynamics track switch dynamics

If a brain region participates in meta-control, then its neural population dynamics should evolve in parallel with the behaviourally derived w_t dynamics[43]. In addition, if a region explicitly represents the current policy state, that information should be linearly accessible from the population activity. We therefore explored whether neural dynamics tracked the temporal evolution of w_t around switch time and whether the current policy value could be decoded from neural activity.

To test this, we aligned the neural activity to target switches, whose trajectory resembles a canonical sigmoidal shape (Figs. 2g and 4a), and decomposed population activity using demixed principal components analysis (dPCA; Fig. 4b and Methods). We then reduced the population activity to a single low-dimensional trajectory by projecting it onto the dPCA component related to w_t. To test if the current policy value could be decoded from neural activity, we projected held-out test trials onto the same w_t component (Methods).

We found that for ACC, the top dPCA component for w_t explained 17.8% of the variance, whereas its interaction with prey value difference explained 22.5% (Fig. 4b). ACC neural dynamics closely mirrored the sigmoidal w_t trajectory (Pearson's $r = 0.99$, $P < 0.0001$). Furthermore, in ACC, switch value could be decoded before and after the switch (decoding accuracies 61.0% and 62.5%, respectively; one-sided permutation test: both $P = 0.03$; Fig. 4c). Together, these results indicate that ACC closely tracks the meta-controller variable throughout the whole switch period and contains an explicit representation of the current policy state at both switch onset and offset.

In HPC, as in ACC, neural population dynamics were explained by the w_t dPCA component (14.2%), and its interaction with differences in prey value (19.9%; Fig. 4b). HPC w_t dPCA projections also resembled the switch trajectories ($r = 0.84$, $P < 0.001$; Fig. 4c), but this correlation was significantly weaker than in ACC (Fisher test for correlation, $P < 0.001$). Moreover, switch value could be decoded from HPC neurons only at switch offset (64% decoding accuracy, $P < 0.001$). These findings suggest that HPC representations primarily reflect the selected policy after the switch, rather than throughout the transition.

In OFC, as in the other areas, neural population dynamics were primarily explained by the w_t dPCA component (14.7%) and its interaction with differences in prey value (20.6%; Fig. 4b). The correlation between the w_t dPCA projections and switch trajectories was comparable to that observed in the HPC ($r = 0.86$, $P < 0.001$; Fig. 4c) and significantly lower than in ACC. However, switch direction could not be decoded from OFC activity at any time point. The absence of significant decoding in OFC may reflect reduced statistical power due to lower neuron counts, but substantial single-neuron tuning to w_t (Fig. 3c) argues that limited sampling alone is unlikely to explain this result. OFC projections converged at switch onset and diverged only after the transition. Thus, OFC may primarily reflect the outcome of the policy transition rather than its ongoing evolution.

The dPCA also indicated substantial variance explained by condition-independent modes in all brain areas (ACC: 12%; HPC: 14.1%; OFC: 18.8%; Fig. 4d). The condition-independent modes represent w_t evolution over time, and although the origins of their dynamics are unclear, visual inspection suggests distinct temporal patterns across brain regions.

Next, we examined how switch representations differ between the equal-value and unequal-value conditions. We re-applied dPCA, this time configuring the decomposition to isolate w_t dynamics across five discrete w_t levels (Fig. 4e and Methods). Across all regions, most variance was explained by the time-dependent w_t component (Fig. 4e). Projecting population activity onto the top dPCA component related to w_t revealed differences in how policy dynamics were structured by reward value (Fig. 4f). In ACC, unequal-value trials produced an asymmetric organization across w_t bins that became more symmetric when values

Table 2 | Mixed-effects model results for the equal-value condition (Δ prey value=0), using the same analyses and conventions as in Table 1

Δ prey value=0	Predictors significance (*P* values for full model)								
Brain region	**Time from switch onset (ms)**	**Time until switch**	**Time elapsed**	**Away distance**	**To distance**	**Away slope**	**LR χ^2**	**LR (*P* value)**	**ΔAIC**
ACC	−400	0.41	0.86	0.85	0.75	0.90	2.17	0.14	0.17
ACC	−100	0.02	0.08	0.54	0.78	0.49	5.84	0.02	3.84

were equal, consistent with the FI results showing value-biased blending policies (Figs. 3g and 4f). In HPC, activity symmetrically spanned the full w_t continuum in the different-value condition, but shifted toward blended states when values were equal (Fig. 4f). OFC projections were largely symmetric across reward conditions (Fig. 4f).

Correspondingly, ACC correlations with HPC and OFC were reduced in the equal-value condition, while HPC–OFC correlations increased (Fisher test, $P < 0.001$; Fig. 4g,h), indicating that reward structure reshaped how policy representations were coordinated across brain regions.

Neural ramping during goal switching

In discrete choices, ACC population activity often exhibits ramping dynamics aligned to decisions[36,44]. We hypothesized that ACC may have an analogous role in continuous choices (Fig. 5a).

We used a method that combines *k*-means clustering with linear dynamic systems, which enabled us to cluster neurons based on their dynamics (Methods). The number of clusters was optimized (silhouette score) separately for each brain region and reward condition, and cluster selectivity for target switches was then assessed using two-sided permutation tests comparing firing rates during switch and control (non-switch) windows. Only clusters showing significant differences were retained.

In ACC, this process yielded three ACC clusters (unequal-value condition, permutation test, $P < 0.01$; Fig. 5a,b). The largest cluster (cluster 1, $n = 45$ neurons) exhibited ramping activity that peaked at switch time and spanned the entire switch window (Fig. 5a, dark green trace). Ramping associated with switches from higher-value to lower-value prey was stronger and peaked later (33 ms). Notably, for both high-to-low and low-to-high switches, neural activity preceded the switch time (−150 ms; Fig. 4a), suggesting that ACC modulation precedes switching. The sustained activity across the switch window is consistent with a monitoring role for ACC.

In the HPC, we found four clusters of neurons (permutation tests, $P < 0.01$; Fig. 5c,d). Cluster 1 qualitatively resembled the ACC ramping signal, although it differed quantitatively: it started at a higher baseline (ACC mean *z*-scored firing rate at switch start = −0.64; HPC = −0.37; two-sided *t*-test: $P < 0.001$) and reached a lower peak (ACC at switch time = 0.46; HPC = 0.33; two-sided *t*-test: $P < 0.001$), resulting in a reduced slope (ACC mean slope: 1.09; HPC: 0.70; two-sided *t*-test: $P < 0.001$). Neurons in cluster 2, instead, tracked the w_t over time (Fig. 5c), consistent with a representation of policy state rather than switch-related ramping (see also Fig. 4a). These responses differed across switch directions, indicating that some HPC neurons are sensitive to value. Clusters 3 and 4 were mostly overlapping across switch directions, indicating that some neurons are not sensitive to value (for further characterization, see Supplementary Note 2). Moreover, neurons in cluster 3 recapitulated the dynamics observed in population-level decoding (Fig. 4c), exhibiting increased activity after the switch rather than ramping toward the decision point ($t = 0$ ms).

In OFC, we identified three clusters (all permutation tests, $P < 0.01$; Fig. 5e,f), each exhibiting distinct temporal dynamics, without clear ramping toward the switch in any of them.

Overall, unequal-value trials revealed distinct switch-related dynamics across brain regions: ACC showed sustained ramping that peaked at switch time, HPC showed a weaker ramp together with mixed value-sensitive and value-insensitive dynamics, and OFC exhibited heterogeneous non-ramping profiles. Interestingly, one OFC cluster resembled the policy-like dynamics observed in HPC cluster 2, suggesting a partial convergence in how these regions represent ongoing policy states despite otherwise distinct temporal organizations.

In the equal-value condition trials, ACC had three clusters ($P < 0.01$; Fig. 5g); cluster 1 recapitulated the ramping observed in the different-value condition, demonstrating that ramping is not an artefact of value difference, but likely relates to the decision to switch. Cluster 2 exhibited a temporal profile reminiscent of the behavioural policy; cluster 3 did not map onto any recognizable switch-related signature. In HPC, we found two significant clusters ($P < 0.01$; Fig. 5h), both of which exhibited multiphasic ascending and descending temporal profiles reminiscent of the sigmoidal shape of the behavioural policy, even when value asymmetries are removed. In OFC, we only found a single significant cluster showing an ascending temporal profile that peaked at the end of the switch ($P < 0.01$; Fig. 5i).

Distinct ramping in ACC and HPC

The ramping observed in ACC could reflect a prediction of the impending switch, but it could also arise from variation in behavioural variables that are themselves predictive of switching (Fig. 2h and Extended Data Fig. 7). To isolate switch activity, we regressed out these behavioural covariates and looked at the residual firing, which revealed a ramp emerging immediately (−100 ms) before the switch (Fig. 5j). We then quantified this effect by fitting linear mixed-effects models to ramping neurons (both reward conditions), including behavioural covariates and a time-until-switch regressor (counting down to the upcoming switch), across 400 ms and 100 ms pre-switch windows ($t = 0$ ms in Figs. 4a and 5a and Methods). In ACC, time until switch explained firing beyond behavioural covariates: in the different-reward condition, the full model outperformed the reduced model in both windows (Table 1), whereas in the equal-reward condition this effect remained significant in the immediate pre-switch window (Table 2).

We then repeated this analysis in cluster 1 HPC neurons. The full model outperformed the reduced model only in the 400 ms window before the switch, with a smaller difference in Akaike information criterion (ΔAIC) than ACC (Table 1), but this effect did not persist in the immediate pre-switch period. Accordingly, time-until-switch did not explain HPC firing in this late window, while behavioural covariates continued to contribute. Residual activity did not show a late increase approaching the switch, but instead exhibited an early peak followed by a relatively flat trajectory (Fig. 5k).

Together, these findings indicate that ACC carries sustained predictive signals of impending switches, whereas HPC activity is more consistent with an earlier, transient planning-related signal that precedes but does not extend into switch execution.

Discussion

Behaviour in naturalistic, continuous, interactive contexts is driven by goals, which can be blended to produce a composite behaviour. Here we use the idea of compositional control to model behaviour and its neural foundations in a continuous multi-prey-pursuit task.

Our control-theoretic decomposition of behaviour resembles ethogramming[45,46]. However, whereas standard ethogramming approaches assume that behavioural states are discrete and exclusive, our method estimates blended strategies.

Our neural results challenge the traditional division of labour between the medial temporal lobe and the frontal executive system, raising the possibility of the HPC as a central node in the real-time control of complex, multi-variable behaviour. Some HPC neurons (cluster 1) showed a transient component that was related to anticipatory variables over a broader time window preceding the switch. However, these signals did not persist at the time of commitment, pointing to an early, transient representation of impending switches rather than a sustained predictive signal[40]. Our results do not, however, imply that other regions have no role in continuous choice. For example, parietal cortex and dorsolateral prefrontal cortices almost certainly have major roles[47].

Recent theorizing on the HPC emphasizes its role in representing relational structures[20,21,48]. The cognitive map theory posits that the HPC creates structured representations of the world by composing and binding individual basis elements into a cognitive map[18,21,49]. Our neural results align with and extend the cognitive map theory to the domain of control. That is, we found that the HPC also encodes compositions that bind goal-specific control policies.

Indeed, our findings suggest an active role for HPC in choice: as a state-estimator controller for dynamic behaviour. That is, a system that represents current latent states, plans future states and updates the representations through feedback[50]. Although the set of studies positing potential control functions remains relatively small, several recent works show HPC functions integrating with decision-making and planning[51–54].

The present results indicate that ACC carries a robust meta-control signal. This potential control-related role relates to a long tradition that emphasizes the role of ACC in control[35,55–57] and meta-control. The ramping effects, in turn, are reminiscent of an extensive literature demonstrating that ACC accumulates evidence for strategic shifts in decision-making or rule application[31,36,44]. Notably, meta-control can be either implicit or explicit; our data, which show clear explicit monitoring and meta-control signals in ACC, provide evidence for the explicit implementation. We also find that this ramping switch signal is dependent on the value of switching goals, a finding in line with previous studies and theories linking ACC to integrating cost-benefit information into decisions[33,58,59]. These results clearly implicate ACC in the meta-control aspect in a way that HPC and OFC are not as clearly involved.

Our results suggest that OFC has a minimal role in policy switching. Value is an important input for control[60]. A value-input role would be consistent with known economic roles for OFC[37,38]. OFC may also organize information into a cognitive map of task space[26,39]. Furthermore, our data show that OFC activity evolved with switching, but unlike the other brain regions, OFC representations converged at switch onset and diverged only after the transition, suggesting that OFC contextualizes the selected policy after the switch rather than tracking its ongoing evolution.

Ultimately, our results take on classic questions of neuroeconomics regarding how we represent the space of possible choices, assign values to them and adjudicate between them, but extend them to the continuous domain. Doing so reveals the fundamental unification between choice and control; moreover, it argues for control theory, rather than microeconomic theory, as an intellectual foundation for a neuroscience of continuous choice[3].

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/s41586-026-10896-8.

1. Gordon, J. et al. The road towards understanding embodied decisions. *Neurosci. Biobehav. Rev.* **131**, 722–736 (2021).
2. Maselli, A. et al. Beyond simple laboratory studies: developing sophisticated models to study rich behavior. *Phys. Life Rev.* **46**, 220–244 (2023).
3. Yoo, S. B. M., Hayden, B. Y. & Pearson, J. M. Continuous decisions. *Philos. Trans. R. Soc. B* **376**, 20190664 (2021).
4. Merel, J., Botvinick, M. & Wayne, G. Hierarchical motor control in mammals and machines. *Nat. Commun.* **10**, 5489 (2019).
5. Cisek, P. Making decisions through a distributed consensus. *Curr. Opin. Neurobiol.* **22**, 927–936 (2012).
6. Gallivan, J. P., Chapman, C. S., Wolpert, D. M. & Flanagan, J. R. Decision-making in sensorimotor control. *Nat. Rev. Neurosci.* **19**, 519–534 (2018).
7. Yoo, S. B. M., Tu, J. C., Piantadosi, S. T. & Hayden, B. Y. The neural basis of predictive pursuit. *Nat. Neurosci.* **23**, 252–259 (2020).
8. Fabian, S. T., Sumner, M. E., Wardill, T. J., Rossoni, S. & Gonzalez-Bellido, P. T. Interception by two predatory fly species is explained by a proportional navigation feedback controller. *J. R. Soc. Interface.* **15**, 20180466 (2018).
9. Sridhar, V. H. et al. The geometry of decision-making in individuals and collectives. *Proc. Natl Acad. Sci. USA* **118**, e2102157118 (2021).
10. Yang, Q. et al. Monkey plays Pac-Man with compositional strategies and hierarchical decision-making. *eLife* **11**, e74500 (2022).
11. Bertsekas, D. P. in *Encyclopedia of Optimization* (eds Pardalos, P. M. & Prokopyev, O. A.) 1–6 (Springer, 2025).
12. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
13. Theodorou, E. A., Buchli, J. & Schaal, S. A generalized path integral control approach to reinforcement learning. *J. Mach. Learn. Res.* **11**, 3137–3181 (2010).
14. Dvijotham, K. & Todorov, E. in *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control* (eds Lewis. F. L. & Liu, D.) 119–141 (Wiley, 2012).
15. Wolpert, D. M. & Kawato, M. Multiple paired forward and inverse models for motor control. *Neural Netw.* **11**, 1317–1329 (1998).
16. Lake, B. & Baroni, M. Generalization without systematicity: on the compositional skills of sequence-to-sequence recurrent networks. In *Proc. 35th International Conference on Machine Learning* (eds. Dy, J. & Krause, A.) 2873–2882 (PMLR, 2018).
17. Todorov, E. Compositionality of optimal control laws. In *Advances in Neural Information Processing Systems 22* (eds Bengio, Y. et al.) (NeurIPS, 2009).
18. Kurth-Nelson, Z. et al. Replay and compositional computation. *Neuron* **111**, 454–469 (2023).
19. Whittington, J. C. R., McCaffary, D., Bakermans, J. J. W. & Behrens, T. E. J. How to build a cognitive map. *Nat. Neurosci.* **25**, 1257–1272 (2022).
20. Eichenbaum, H. & Cohen, N. J. Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron* **83**, 764–770 (2014).
21. Behrens, T. E. J. et al. What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* **100**, 490–509 (2018).
22. Kay, K. et al. Constant sub-second cycling between representations of possible futures in the hippocampus. *Cell* **180**, 552–567.e25 (2020).
23. Park, S. A., Miller, D. S., Nili, H., Ranganath, C. & Boorman, E. D. Map making: constructing, combining, and inferring on abstract cognitive maps. *Neuron* **107**, 1226–1238.e8 (2020).
24. Sanders, H., Wilson, M. A. & Gershman, S. J. Hippocampal remapping as hidden state inference. *eLife* **9**, e51140 (2020).
25. Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054–1069 (2011).
26. Wikenheiser, A. M. & Schoenbaum, G. Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nat. Rev. Neurosci.* **17**, 513–523 (2016).
27. Yeung, N. & Summerfield, C. Metacognition in human decision-making: confidence and error monitoring. *Philos. Trans. R. Soc. B* **367**, 1310–1321 (2012).
28. Eppinger, B., Goschke, T. & Musslick, S. Meta-control: From psychology to computational neuroscience. *Cogn. Affect. Behav. Neurosci.* **21**, 447–452 (2021).
29. Musslick, S., Cohen, J. D. & Goschke, T. in *Encyclopedia of the Human Brain* (ed. Grafman, J. H.) 269–285 (Elsevier, 2025).
30. Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* **14**, 1338–1344 (2011).
31. Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J. & Rushworth, M. F. S. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).
32. Kolling, N., Behrens, T., Wittmann, M. & Rushworth, M. Multiple signals in anterior cingulate cortex. *Curr. Opin. Neurobiol.* **37**, 36–43 (2016).
33. Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013).
34. Akam, T. et al. The anterior cingulate cortex predicts future states to mediate model-based action selection. *Neuron* **109**, 149–163.e7 (2021).
35. Heilbronner, S. R. & Hayden, B. Y. Dorsal anterior cingulate cortex: a bottom-up view. *Annu. Rev. Neurosci.* **39**, 149–170 (2016).
36. Sarafyazd, M. & Jazayeri, M. Hierarchical reasoning by neural circuits in the frontal cortex. *Science* **364**, eaav8911 (2019).
37. Padoa-Schioppa, C. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* **34**, 333–359 (2011).
38. Hunt, L. T. et al. Triple dissociation of attention and decision computations across prefrontal cortex. *Nat. Neurosci.* **21**, 1471–1481 (2018).
39. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279 (2014).
40. Elston, T. W. & Wallis, J. D. Context-dependent decision-making in the primate hippocampal–prefrontal circuit. *Nat. Neurosci.* **28**, 374–382 (2025).

41. Yoo, S. B. M., Tu, J. C. & Hayden, B. Y. Multicentric tracking of multiple agents by anterior cingulate cortex during pursuit and evasion. *Nat. Commun.* **12**, 1985 (2021).
42. Wilson, R. C. & Collins, A. G. E. Ten simple rules for the computational modeling of behavioral data. *eLife* **8**, e49547 (2019).
43. Khona, M. & Fiete, I. R. Attractor and integrator networks in the brain. *Nat. Rev. Neurosci.* **23**, 744–766 (2022).
44. Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939 (2011).
45. Anderson, D. J. & Perona, P. Toward a science of computational ethology. *Neuron* **84**, 18–31 (2014).
46. Brown, A. E. X. & de Bivort, B. Ethology as a physical science. *Nat. Phys.* **14**, 653–657 (2018).
47. Milner, D. & Goodale, M. *The Visual Brain in Action* (Oxford Univ. Press, 2006).
48. Gershman, S. J. & Niv, Y. Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* **20**, 251–256 (2010).
49. Zutshi, I. et al. Hippocampal neuronal activity is aligned with action plans. *Nature* **639**, 153–161 (2025).
50. Shadmehr, R. & Krakauer, J. W. A computational neuroanatomy for motor control. *Exp. Brain Res.* **185**, 359–381 (2008).
51. Bakkour, A. et al. The hippocampus supports deliberation during value-based decisions. *eLife* **8**, e46080 (2019).
52. Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive map. *Nat. Neurosci.* **20**, 1643–1653 (2017).
53. Vikbladh, O. M. et al. Hippocampal contributions to model-based planning and spatial memory. *Neuron* **102**, 683–693.e4 (2019).
54. Edelson, M. G. & Hare, T. A. Goal-dependent hippocampal representations facilitate self-control. *J. Neurosci.* **43**, 7822–7830 (2023).
55. Cohen, J. D., Botvinick, M. & Carter, C. S. Anterior cingulate and prefrontal cortex: who's in control?. *Nat. Neurosci.* **3**, 421–423 (2000).
56. Verguts, T. Binding by random bursts: a computational model of cognitive control. *J. Cogn. Neurosci.* **29**, 1103–1118 (2017).
57. Silvetti, M., Vassena, E., Abrahamse, E. & Verguts, T. Dorsal anterior cingulate-brainstem ensemble as a reinforcement meta-learner. *PLoS Comput. Biol.* **14**, e1006370 (2018).
58. Friedman, A. et al. A corticostriatal path targeting striosomes controls decision-making under conflict. *Cell* **161**, 1320–1333 (2015).
59. Kane, G. A. et al. Rat anterior cingulate cortex continuously signals decision variables in a patch foraging task. *J. Neurosci.* **42**, 5730–5744 (2022).
60. Botvinick, M. M., Niv, Y. & Barto, A. G. Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition* **113**, 262–280 (2009).

Methods

Human intracranial neurophysiology

Experimental data were recorded from 19 adult patients (8 men and 11 women; mean age 40.1 ± 13.1 years; Extended Data Tables 1 and 2) undergoing intracranial monitoring for epilepsy (for general recording methods, see ref. 61). The HPC, ACC and OFC were not a seizure focus area of any patients included in the study. Single-neuron data were recorded from stereotactic (sEEG) probes, specifically AdTech Medical probes in a Behnke-Fried configuration. Each patient had an average of four probes. Each probe includes eight microwires (that is, eight contacts), specifically designed for recording single-neuron activity. Single-neuron data were recorded using a 512-channel Blackrock Microsystems Neuroport system sampled at 30 kHz. To identify single-neuron action potentials, the raw traces were spike-sorted using the WaveClus[62] sorting algorithm and then manually evaluated. Noise was removed and each signal was classified as multi or single unit using several criteria: consistent spike waveforms, waveform shape (slope, amplitude, trough-to-peak), and exponentially decaying inter-spike interval histogram, with no inter-spike interval shorter than the refractory period (1 ms). We used only single-neuron activity for all analyses. The experiment's procedures were conducted by the standards set by the Declaration of Helsinki and approved by the Institutional Review Board for Baylor College of Medicine, Houston, TX (H-18112). Participants provided written informed consent after the experimental procedure had been fully explained and were reminded of their right to withdraw at any time during the study. Participants included in our study received a standard monetary compensation, which did not depend on task performance.

Electrode visualization

Electrode locations were localized using the intracranial electrode visualization pipeline (iELVis[63]). Postoperative CT scans were co-registered to preoperative MRI space using FSL[64,65], and electrode contacts were manually identified in BioImage Suite (v3.5[66]). Electrode coordinates were transformed to native anatomical space and visualized on Freesurfer-reconstructed cortical surfaces (v7.4.1[67]) using iELVis[68]. Microelectrode locations were assigned to the deepest macro contact of the AdTech Behnke-Fried electrodes. For group visualization, electrode coordinates were transformed into MNI152 space and plotted on a glass brain using Reproducible Analysis and Visualization of intracranial electroencephalography (RAVE[69,70]; Fig. 3a).

Healthy participants

In addition to the dataset with intracranial recordings, behavioural data were collected from 50 healthy adult participants who performed the same joystick-controlled prey-pursuit task but without physiological recordings (Extended Data Table 1). Of these participants, 33 were recruited at Baylor College of Medicine, and the remaining 17 were recruited at the University of Minnesota. The experiment's procedures were conducted by the standards set by the Declaration of Helsinki and approved by the Institutional Review Board for Baylor College of Medicine, Houston, TX (H-54759) and University of Minnesota (00009204). Participants provided written informed consent after the experimental procedure had been fully explained and were reminded of their right to withdraw at any time during the study. Participants included in our study received a standard monetary compensation, which did not depend on task performance.

Experimental apparatus

To control their computer avatar, participants used a joystick that was a modified version of a commercially available joystick with a built-in potentiometer (Logitech Extreme Pro 3D). The joystick position was read out in the MATLAB psychophysics toolbox running on the stimulus control computer.

Task description

All participants performed the prey-pursuit task[7,41,71] (see Supplementary Video 1). At trial start, two or three shapes appeared on a grey computer monitor placed directly in front of the participant. The yellow circle (15 pixels in diameter) was an avatar that represented the participant and began at the centre of the screen. Participant position was determined by the joystick and was limited by the screen boundaries. A square (30 pixels in length) represented the prey(s). Prey movement was determined by a simple algorithm (see below). Each trial ended with either the successful capture of the prey or after 20 s, whichever came first. Successful capture was defined as any spatial tangency between the avatar circle and the prey square. Capture resulted in an immediate numeric reward. Prey movement was generated interactively using an A-star algorithm. Specifically, for every frame (16.67 ms), we computed the cost of 15 possible future positions the prey could move to in the next time-step. These 15 positions were equally spaced on the circumference of a circle centred on the current position of the prey, with a radius equal to the maximum distance the prey could travel within one time-step. The cost was based on two factors: the position in the field and the position of the participant's avatar. The field that the prey moved in had a built-in bias for cost, which made the prey more likely to move toward the centre. The cost due to distance from the avatar of the participant was transformed using a sigmoidal function: the cost became zero beyond a certain distance so that the prey did not move, and it became greater as distance from the avatar of the participant decreased. From these 15 positional costs, the position with the lowest cost was selected for the next movement. If the next movement was beyond the screen range, then the position with the next lowest cost was selected until it was within the screen range. The maximum participant speed was 23 pixels per frame (each frame lasted 16.67 ms). The maximum and minimum speeds of the prey varied across individuals and were set by the experimenter to obtain a large number of trials. Specifically, speeds were selected so that participants could capture prey on no more than 85% of trials. To ensure sufficient time of pursuit, the minimum distance between the initial position of each individual avatar and prey was 400 pixels.

Data structure

The study involved repeated measurements within participants, including multiple trials per participant and multiple neurons recorded from each participant.

Behavioural modelling

Overview of control choice decomposition. The problem we are studying here is how individuals decide between how to pursue goals on a moment-to-moment basis. This means we must estimate a decision variable over how targets are selected to drive control actions. We motivate this problem by first considering the low-level, individual goal problem.

For a given target, participants must deploy some policy that dictates their actions for reaching that goal, $u_k(t)$. Here we consider the controls as the output of a linear controller that is driven by a set of gains that dictate controls (library of controller types considered in Extended Data Table 3). For example, the simplest proportional controller with a gain (g_k) that minimizes distance to target (e_k) for a single target k is:

$$u_k(t) = g_k e_k(t)$$

When there are multiple goals, and each target has different controllers gains, inferring a decision variable requires inferring how a weighted mixture of targets drives participants' control actions[28,72]. To do this, we used ideas from control theory to develop a method for decomposing the participants' instantaneous control (that is,

acceleration) into distinct contributions from each target. This previous work has shown control in a nonlinear cost landscape—when there are multiple targets and thus minima in the landscape—can be approximated by a compositional mixture of linear controllers, and the mixing weights in optimal settings are given by the relative cost/value of individual controllers[14,17]. Next, we state the general problem and connection to common mixture models, then our approach to estimating this quantity from behavioural data[73,74].

At every time point, we want to decompose the individual's measured control signal ($u_{obs}(t)$ = acceleration) into the proportion of their control signal attributable to individual prey (k) targets ($u_k(t)$).

$$p(u_{obs}(t)) = \sum_{i=1}^{K} p(\text{target}_k | u_{obs}(t))\ u_k(t)$$

These controls are weighted by the posterior probability of them following a target, given their control signal. Estimation of these weights implicitly defines a multinomial posterior probability over K targets given the control signal. This posterior is the same form as the familiar Gaussian mixture model[75]. Posterior latent target probabilities are then computed via Bayes rule:

$$w_k(t) \propto p(\text{target}_k | u_{obs}(t)) = \frac{p(u_{obs}(t)|\text{target}_{k=1})p(\text{target}_{k=1})}{\sum_{k=1}^{K} p(u_{obs}(t)|\text{target}_{j=1})p(\text{target}_{j=1})}$$

This shows that estimating $w_k(t)$ will implicitly define the integration of an observed data likelihood and categorical prior. Because solving for the weights is a convex optimization problem, we will explicitly optimize (see Model parameterization) the controller weights $w_k(t)$:

$$p(u_{obs}(t)) = \sum_{i=1}^{K} w_k(t) u_k(t)$$

A final note worth mention is how optimizing $w_k(t)$ (that is, target posterior) as a mixture model mixture approach is different from hidden Markov models. Mixture modelling applies soft-assignment or weighting and, as such, using $w_k(t)$ to explain participants' control signal allows blending of controllers and smooth transitions between them. Hidden Markov models use state decoding or hard-assignment, and effectively apply an argmax[$w_k(t)$] to choose a singular controller. In the supplementary (section: model optimization) we show with simulations using a soft-assignment picks up both cases, where switching is either gradual or binary between targets; thus, using soft-assigned weights will correctly detect the target pursuit and switching strategy for either blending of controllers (soft assignment) or discrete controller choices (hard assignment).

Model parameterization. To infer an estimated $w_k(t)$ time series per trial, we estimated maximum a posteriori parameters from behavioural data, including the parameters of $w_k([\phi,\psi])$ and controller gains ($g_k = \theta$). Both sets of parameters are explained below. Optimization of these parameters proceeded by minimizing the negative log posterior:

$$-\log p(u_{obs}) \propto \sum_{t=1}^{T} -\log p(u_{\text{pred},K}(t;\theta,\phi,\psi)) - \log p(\theta,\phi,\psi)$$

The first term on the right side of the above equation is the negative log likelihood (priors are second). u_{pred} is the model predicted control signal that is determined by the mixture of controls across targets:

$$u_{pred}(\theta,\phi,\psi) = w_1(t;\phi,\psi)u_{\text{pred},k=1}(t;\theta) + w_2(t;\phi,\psi)u_{\text{pred},k=2}(t;\theta)$$

Parameterization of w_k. Mixture weights (w_k) were parameterized by implicitly modelling time-dependent logits with ($n = 30$) radial basis functions equally-spaced over each trial's time. Basis parameters included weights (ϕ) and a shared width parameter (ψ). We chose to implicitly parameterize w_k with basis functions as it allows a flexible and smooth parameterization of transitions between target control blending. Practically, w_k values are defined though a softmax over the logits. Logits are computed by multiplying basis functions by their optimized parameters.

$$w_1(t) = \text{softmax}(\text{logits}_1(t;\phi,\psi))$$

$$w_2(t) = 1 - w_1(t)$$

Note, this approach is readily extended to the case of $K > 2$ targets by constraining one target's logits to 1 and increasing the basis parameters to $nK - 1$.

Parameterization of $u_{pred,k}$. For generating predicted controls per target, $u_{pred,k}$, we tested different controller types and estimate their gain parameters per controller (g_k). We model the predicted subject's states (x_{pred} = {positions, velocities}) and controls as a linear dynamic system mapping joystick input to the cursor. Joystick position controlled the cursor velocity, and therefore control signals from the subject (u_{obs}) are acceleration, which are composed of the controls from distinct targets ($u_{pred,k}$). These controls and states step the predictions forward and generate a new set of control predictions in each optimization iteration. The linear system describing the subject's cursor is:

$$x_{pred}(t+1) = Ax_{pred}(t) + Bu_{pred}(t)$$

Where A and B are known system parameters (Δt =sampling rate (1/60 Hz)) that determine the mapping from joystick to subject cursor dynamics.

$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & \Delta t \end{bmatrix}$$

Because the subject's control strategy for a given target is unknown, we must also fit the mixing weights using different combinations of controller classes that could generate responses. We considered controllers that use different combinations of error signals to drive target pursuit. Model classes included either proportional or integral errors, where the error signals were of different classes described below, including feedback and feedforward predictive controller errors. All model classes are listed in Extended Data Table 3.

Note that we also included models with future position errors. This controller class allows the subject's predicted controls to be driven by feedforward signals based on the target's predicted (one step ahead) position. The error signal encodes the difference between the subject's current (t) position and target's future ($t+1$) position. Target future position was computed using the forward kinematics of the target:

$$x_{target}(t+1) = x_{target}(t) + \Delta t x'_{target} + \Delta t^2 x''_{target}$$

Model loss priors. The parameter priors for the loss are denoted as:

$$p(\theta)p(\psi)p(\phi) = N(0, I\sigma_\theta^2)N(0, \sigma_\psi^2)N(0, (\lambda S)^{-1})$$

The prior over controller gains (θ) was a zero-mean normal distribution, with identity variance-constraining gains independently. A similar prior was used for the basis function width (ψ). For both priors, increasing σ^2 decreases the regularization of the prior. For basis function weights (ϕ), we wanted to penalize deviations from smoothness. For this, we used a Gaussian Laplacian prior (multivariate Gaussian)

often used in spline models to penalize the wiggliness of functions[76]. This prior uses the inverse precision matrix with a regularization coefficient ($(\lambda S)^{-1}$). S is a penalty matrix defined as the second-order difference between adjacent weights, penalizing the curvature:

$$\Delta^2\phi = \phi_{i+1} - 2\phi_i + \phi_{i-1}$$

Model optimization. For parameter estimation, we used trust-const in scipy.optimize. During each parameter update, the current parameter set (θ,ϕ,ψ) is used to simulate the predicted dynamics and compute the loss. Loss Jacobians and Hessians were computed automatically using Jax and parameters were iteratively optimized by maximizing the negative log posterior. To ensure positively constrained controller gains (θ) and basis width (ψ), we optimized their log and used a softplus transform. Each trial was optimized separately. For each optimization run, we fit a specific controller class (same for both targets). This was repeated four times for each controller class, with different parameter initializations, and the one with best loss was retained for cross model performance comparisons (next section). This was repeated for each of the seven model classes (Extended Data Table 3).

Quantifying model performance and computing mixture. After optimizing parameters and retaining the best fit for each model class (and each trial), we computed the evidence lower bound (ELBO) for each model using a Laplace approximation to the parameter posteriors[75]. The ELBO provides an estimate of lower bound on the marginal log likelihood of each model, balancing model fit and penalizing the number of parameters.

The next step is computing a mixing trajectory (for example, w_1) for each trial that marginalizes over each model, while accounting for each model's uncertainty in predicted parameters. Because the model probability is quantified via the ELBO, we can use these ELBOs to compute posterior probabilities for each model.

$$p(M_k|u_{\text{obs}}) = \frac{\exp(ELBO_m)}{\sum_1^j \exp(ELBO_j)}$$

We use these model probabilities to perform Bayesian model averaging (BMA) to obtain a singular w_1 trajectory. In short, the w_1 trajectory around the maximum a posteriori parameters for each model class is weighted and summed. Computing w_1 through model averaging accounts for model uncertainty by marginalizing over the generating models, and this is the target blending estimate (BMA) that we use for downstream analysis.

$$\text{BMA}(w_1) = w_{1,k}p(M_k|u_{\text{obs}})$$

Comparison of different control mixing strategies

We also want to compare a smooth controller blending model to two other possible controller selection models that could drive behaviour. We can use the same model fitting framework as above. For one comparison model, we fit a model that assumes subject's choose a single controller at an given moment. This is equivalent to argmax model over individual target controllers, is nested within the smooth blending model, and resembles how an HMM would choose the generative controller. Comparison of this model to the smooth blending effectively tells us how much of behaviour is explainable by a continuous mixing versus discrete selection strategy. This argmax model is formulated for the predicted control as:

$$u_{\text{pred}}(\theta,\phi,\psi) = \text{argmax}_{u_{\text{pred},k}}(w_1(t;\phi,\psi), w_2(t;\phi,\psi))$$

Another strategy that subjects could use is blending inputs with a single controller. Note that this control strategy is not nested within the smooth controller blending model, as control signals produced by multiple controllers versus one produced by weighting multiple inputs are not related through any linear transform (Supplementary Note 1). The input blending model is formulated using a single set of controller gains rather than distinct gains per target, and weighting the inputs:

$$u_{\text{obs}}(t) = \theta \sum_{i=1}^{K} w_k(t)\, x_k(t)$$

Both comparison models were fit using the same approach as the controller blending model. Bayesian model selection with the ELBO was used to select the best-fitting model for each trial. For each subject, we computed the mean posterior probability of each model type for group level model selection.

Behavioural analysis

Finding behavioural switch points. To separate instantaneous moments into pursuit versus change-of-mind points between targets, we first searched each trial to find points where $w_t = 0.5$. These points denote crossings between targets. (There is necessarily some element of estimation in the process of detecting switches, and, due to the inherent variability in human behaviour, our algorithm necessarily has, at some point, an arbitrary cutoff.) We excluded any putative switch points that occurred within 325 ms of the trial start time, under the assumption that these likely represent a correction to an initial choice rather than true pursuit switching or change of mind. Using the remaining potential switch points, we created a window of −250 to +250 ms around these points. The switch trajectories were then categorized as falling into one of three categories. The main category (1) was those where a full switch was made between targets, the other two were either (2) partial switches or (3) unclassifiable. To achieve this classification, we used a PCA across all switch trajectories from within a session. The input matrix for the PCA had switches index as rows and samples as columns (switches × samples). The first principal component always captured the sigmoidal shaped full switch trajectory of interest. We computed the cosine similarity of this first principal component score and all switch trajectories and retained those with a similarity of >0.9. We ran a PCA again on this subset, with the first principal component again producing the canonical sigmoidal trajectory. In this case, we retained those with cosine similarity of >0.97. This threshold proved to be conservative enough to remove all partial switches, based on visual inspection. In total, an average of 35% of crossing-points were labelled as legitimate switches and forwarded for further analysis. Switch onset and offset timing were determined using a linear fitting change-point algorithm that looked for changes in mean and slope of the trajectory. To find the switch start, we searched backward in time from the crossing of $w_t = 0.5$. Switch ends were found by searching forward in time. Note that if two or more switches occur in the defined time window, their combined trajectory no longer resembles a clean sigmoidal transition. Because our PCA based classification relies on cosine similarity to the first principal component (which captures the prototypical sigmoid shape), such multi switches events would yield low similarity scores and would not be labelled as 'full switches'. In our exploration of the data, we find that such double switches are very rare, so our approach is fundamentally conservative.

Switch statistics. The factors underlying the propensity to make target switches were modelled using a Bayesian binomial mixed-effects model implemented in statsmodels in python[77]. We used this approach to allow modelling all participants together, while accounting for individual variation in predictors. Predictor significance was determined using the 95% posterior credible interval. In terms of predictors, we asked if switch probability was sensitive to the time elapsed in trial, the difference in prey reward, and several prey distance variables. Specifically, we included the mean distance to the prey being pursued

and the mean distance of the prey switched to, where the mean is calculated in the 100 ms leading up to a switch. We also included a factor to determine the mean rate of change of distance (distance closure) between both preys, too. The distance closure captures whether either prey getting closer or farther away (on average) altered switch probability. Each of these variables was centred by using their median values extracted from trials in which there were no switches. This allows interpreting the coefficients in terms of changes with respect to control, no-switch trials.

Neural analysis

Single-neuron encoding of task variables. To assess how neurons are tuned to task variables, we employed a linear-nonlinear Poisson model as a GAM[76,78]. Specifically, we aimed to quantify how neurons were tuned to subject speed, relative prey distance, relative prey speed, relative time in trial, prey reward difference, and the model-based blending variable. As our main goal was estimating whether neurons exhibited disentangled (linear) tuning to w_t or multiplexed with prey value, we included both w_t and its interactions with prey value as predictors. The GAM was fit with a penalized regression smoothing spline model[76]. Penalized smoothing spline GAMs were chosen as they have been shown to offer good recovery of nonlinear neuron tuning, particularly in tasks where tuning is estimated with respect to continuously changing variables (for example, speed[78]). Natural cubic regression splines served as a basis function. Cubic splines are convenient as a smooth interpolant basis because their only hyperparameter is the number of bases[76] (J). The spike counts y_t are fit in a GAM with univariate and interaction effects with basis functions b_{j,x_1} (basis j and variable x_1). As an example, in a two-variable model the firing is modelled as:

$$\log(y_t) = \sum_{j=1}^{J} b_{j,x_1}\beta_{j,x_1} + \sum_{j=1}^{J} b_{j,x_2}\beta_{j,x_2} + \sum_{j=1}^{J} (b_{j,x_1} \otimes b_{j,x_2})\beta_{j,x_{1\otimes 2}}$$

The first two terms are examples of single (marginal) variable tuning effects. The latter component encodes the interaction (nonlinear mixed selectivity) term as a tensor over the two univariate basis functions.

The model was fit in a Bayesian framework using stochastic variational inference in NumPyro. To estimate smooth and regularized tuning functions, we follow previous implementations[76,78,79] and implement two types of regularization. Specifically, regularization was implemented through two types of priors on model coefficients that (1) control for basis wiggliness and enforce smoothness across basis functions, and (2) force sparse sets of bases similar to an L1 regularization. Combining these two effectively performs model selection by both smoothing across bases (1) and (2), forcing bases that do not contribute to the model likelihood to be zero. All marginal and interaction terms were included in a single model and subject to the same smoothness and shrinkage priors. As a result, interaction terms that were not supported by the data were automatically shrunk toward zero, yielding an effective main-effects-only model when appropriate. These priors are implemented for basis coefficients as a multivariate normal distribution, for a given variable as:

$$\beta \sim \mathrm{MVN}(0, \Sigma_\beta)$$

$$\Sigma_\beta = (\lambda_j S + \lambda_* S_*)$$

The smoothing/regularization terms (λ) are treated as random variables in the Bayesian approach, with Half-Cauchy priors. This allowed estimating separate smoothing terms per univariate and tensor model term directly from the data, without cross-validation. S encodes a second-order difference penalty between basis coefficients, inducing smoothing as a Laplacian graph prior that is used to construct the coefficients prior covariance. We point the reader to ref. 79 for more specifics on how S and S_* are constructed.

To fit the models and approximate the posterior distribution of parameters, we employed a variational inference approach with a Gaussian guide. Variational parameters (mean and covariance) were optimized by minimizing the negative ELBO. This approach balances computational efficiency with parameter uncertainty estimation. All models were optimized using Adam (learning rate: 10^{-3}), each run for 10,000 steps.

To further determine coefficient significance, 99% posterior credible intervals were used to evaluate whether they included zero. Those variables with intervals non-overlapping with zero were retained. To determine whether neural activity was reliably modulated by task variables, we compared the full model against a baseline intercept-only model using the expected log predictive density (ELPD) estimated via WAIC[80]. The resulting WAIC model weight was used as a neuron-level model selection criterion, and only neurons for which the full model was favoured over baseline were retained for subsequent analyses.

Fisher information of w_t tuning. We computed a cross-validated and denoised estimate of Fisher information using neuron tuning to w_t. We computed tuning curves for each neuron in ten bins. We used fivefold cross-validation when computing the tuning curves. Tuning curves from the training set for each brain region were reduced using principal components analysis, retaining enough dimensions to account for 90% of variance. FI was computed using PCA projections of the test set, and taking the Euclidean distance between adjacent levels of w_t of these projections[81]. To test for significant information, we created a null distribution for permutation testing by permuting w_t labels during calculation of the tuning curves. This distribution was created by repeating this process 1,000 times, and the P value was computed by comparing the mean actual fivefold information to the permutation distribution. Benjamini–Hochberg FDR test was applied to control for false positives. To summarize global differences in tuning, we grouped w_t bins into commitment states (extreme bins: indices 1–2 and 9–10) and blended states (intermediate bins: indices 3–8). FI was averaged within each group, and differences between blended and commitment states, as well as across reward conditions, were assessed using permutation tests.

dPCA. To examine the low-dimensional population dynamics related to behavioural controller blending, we performed dimensionality reduction on neural data using dPCA[82]. Using dPCA, we can extract linear subspaces of population activity that maximize the variance specifically related to separating levels of blending. In effect, this approach allows us to determine whether the population has linearly decodable representations of blending.

Because our aim was assessing whether neural population dynamics closely tracked the behavioural blending variable, we focused on time windows in which there was an explicitly consistent blending trajectory: change of mind between prey targets. To decompose the neural data, we first sorted all switch points, during a trial, for a given subject, and within a session, into those that switched from a low to high or high to low-value target. Under the assumption that blending differs in a mirrored manner (high to low versus low to high w_t), across time, for these switch directions, we can take the decoding subspace as the dynamic subspace for w_t.

We fit the dPCA model to firing rate tensors as pseudo-populations that included all subjects. The model included the effects for w_t, prey reward difference, and their interaction. Each fit was done separately for ACC (n = 254), HPC (n = 546) and OFC (n = 103). The fitting and significance testing procedure was carried out by performing 1,000 repetitions randomly sampling (stratified Monte Carlo) the data and holding out 5% of trials (balanced between switch directions). A dPCA model was then trained using a fixed regularization of 10^{-5}. Firing rate

test sets were then projected along the top decoder dimension for the switch direction. We took the decoded switch direction to be the minimum distance of the projected test data to the projected training data. This classification was computed for each of the 30 time points in the switch window. Second, we established significance thresholds by repeating the above procedure, but randomly swapping trial labels to create a null accuracy distribution. We created the null distribution by repeating the 1,000 iterations 50 times. For each of the 50 sets, we computed the null decoding accuracy. We determined that decoding was statistically significant when the decoding accuracy was greater than the null for 100% of the 50 sets.

To visualize how neural population activity varied across blending states and reward conditions, we performed an additional dPCA analysis in which neural activity was binned by the behavioural blending variable w_t. This analysis was performed separately for trials with equal prey value and trials with different prey value. Neural firing rates were aligned to switch events. For each condition, firing rates were grouped into bins according to the instantaneous value of w_t and averaged to form a population activity matrix spanning time and blending state. We fit the dPCA model to firing rate tensors as pseudo-populations that included all subjects. We applied dPCA with marginalizations over blending state and its interaction with time, and projected population activity onto the leading components associated with these terms. The resulting low-dimensional activity patterns were visualized as heatmaps illustrating how neural dynamics evolved across time and blending state, separately for HPC, ACC and OFC.

Clustering ramp. We clustered firing during the target switches to assess whether classes of single neurons also exhibit ramping behaviour. To obtain a meaningful clustering based on latent neural dynamics and potential, we first fit a linear dynamic system to trial-averaged, Gaussian smoothed (full width at half maximum (FWHM) 60 ms) firing rates using the ssm package in Python. In addition, the models were fit to both switch directions (high-to-low value and low-to-high value prey) simultaneously, giving them shared parameters but different inputs. We used the normalized relative target distance as inputs to the model. This approach to clustering accounts for similar latent dynamics evolving under different inputs. The number of the latent state dynamics for each model were optimized by computing the loss (ELBO) from a held-out test in fivefold cross-validation. This was done separately for each brain region. We found three latent states were suitable for describing each area.

Clustering was then performed on the firing rate embedding matrix C using k-means. The C matrix links latent states to firing rates. The number of clusters was optimized using the silhouette score, separate for each brain region. We chose clusters to retain based on those that exhibited significant differences in activity between time-locked segments with switched and control segments without. Significance was determined by permuting (1,000 times) the mean clustered firing rate between the actual switches and controls.

Mixed-effects model. To quantify if the neural activity carried information about the upcoming switch carried beyond behavioural covariates we used a linear mixed-effects model. For each identified switch and each subject, we aligned neurons firing rates to the time at which w_t crossed 0.5, and constructed a matrix one row per neuron, switch event, and time bin within the pre-switch window. Predictors included behavioural covariates (time elapsed within the trial, Δ prey value, to and away distance, to slope) along with a 'time until switch' predictor. The latter was defined as the time remaining until the switch. All predictors were z-scored prior to model fitting. We assigned a unique label to each neuron and included it as a random intercept in the model. We then fitted a full model including all the predictors and a reduced model lacking the time-until-switch predictor (statsmodels[77] in Python). We obtained the estimates of the coefficients using the restricted maximum likelihood (REML), whereas we compared the models using the likelihood ratio test (LR test) and the Akaike information criterion ($\Delta AIC = AIC_{reduced} - AIC_{full}$), with positive values indicating improved fit of the full model. We repeated the analysis separately for each pre-switch time window (−400 and −100 ms).

Ethics

Experiments were conducted according to protocol guidelines approved by the Institutional Review Board for Baylor College of Medicine and Affiliated Hospitals, Houston TX (H−18112 for the sEEG recordings and H-54759 for the healthy subjects), and by the Review Board for University of Minnesota, Minneapolis, MN (00009204).

Reporting summary

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

Data availability

All preprocessed data are available at https://doi.org/10.6084/m9.figshare.32572764.v3 (ref. 83).

Code availability

All code is available at https://github.com/BCM-Neurosurgery/Prey-Pursuit.

61. Franch, M. et al. A population code for semantics in human hippocampus. Preprint at *bioRxiv* https://doi.org/10.1101/2025.02.21.639601 (2025).
62. Chaure, F. J., Rey, H. G. & Quian Quiroga, R. A novel and fully automatic spike-sorting implementation with variable number of features. *J. Neurophysiol.* **120**, 1859–1871 (2018).
63. Groppe, D. M. et al. iELVis: An open source MATLAB toolbox for localizing and visualizing human intracranial electrode data. *J. Neurosci. Methods* **281**, 40–48 (2017).
64. Jenkinson, M. & Smith, S. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* **5**, 143–156 (2001).
65. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).
66. Joshi, A. et al. Unified framework for development, deployment and robust testing of neuroimaging algorithms. *Neuroinformatics* **9**, 69–84 (2011).
67. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. *Neuroimage* **9**, 179–194 (1999).
68. Yang, A. I. et al. Localization of dense intracranial electrode arrays using magnetic resonance imaging. *Neuroimage* **63**, 157–165 (2012).
69. Magnotti, J. F., Wang, Z. & Beauchamp, M. S. RAVE: Comprehensive open-source software for reproducible analysis and visualization of intracranial EEG data. *Neuroimage* **223**, 117341 (2020).
70. Wang, Z., Magnotti, J. F., Zhang, X. & Beauchamp, M. S. YAEL: your advanced electrode localizer. *eNeuro* **10**, ENEURO.0328-23.2023 (2023).
71. Chericoni, A. et al. Neural geometry in the human hippocampus enables generalization across spatial position and gaze. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2603.04747 (2026).
72. Gómez, V., Kappen, H. J., Peters, J. & Neumann, G. in *Machine Learning and Knowledge Discovery in Databases* (eds Calders, T. et al.) 482–497 (2014).
73. Peng, X. Bin, Chang, M., Zhang, G., Abbeel, P. & Levine, S. MCP: learning composable hierarchical control with multiplicative compositional policies. In *Advances in Neural Information Processing Systems 32* (eds Wallach, H. et al.) (NeurIPS, 2019).
74. Matsuo, Y. et al. Deep learning, reinforcement learning, and world models. *Neural Netw.* **152**, 267–275 (2022).
75. Murphy, K. P. *Probabilistic Machine Learning: An Introduction* (The MIT Press, 2022).
76. Wood, S. N. Generalized additive models. *Annu. Rev. Stat. Appl.* **12**, 497–526 (2025).
77. Seabold, S. & Perktold, J. Statsmodels: econometric and statistical modeling with Python. In *Proc. 9th Python in Science Conference (SciPy 2010)* https://doi.org/10.25080/Majora-92bf1922-011 (SciPy, 2010).
78. Balzani, E., Lakshminarasimhan, K., Angelaki, D. & Savin, C. Efficient estimation of neural tuning during naturalistic behavior. In *Advances in Neural Information Processing Systems 33* (eds. Larochelle, H. et al.) (NeurIPS, 2020).
79. Wood, S. N. *Generalized Additive Models: An Introduction with R* (CRC Press/Taylor & Francis Group, 2017).
80. Gelman, A., Hwang, J. & Vehtari, A. Understanding predictive information criteria for Bayesian models. *Stat. Comput.* **24**, 997–1016 (2014).
81. Kriegeskorte, N. & Wei, X.-X. Neural tuning and representational geometry. *Nat. Rev. Neurosci.* **22**, 703–718 (2021).
82. Kobak, D. et al. Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).

83. Chericoni, A. Dataset for: Neural basis of compositional control. *Figshare* https://doi.org/10.6084/m9.figshare.32572764.v3 (2026).

Acknowledgements We thank P. Schrater, X. Pitkow, J. Pearson, S. Linderman, H. Ritz and A. Shenhav for invaluable discussions and insights on solving the controller problems; J. Johnston for discussion on neural analysis; J. Adkinson, R. Mathura and V. Pirtle for invaluable assistance; and A. Thome for early and visionary efforts in developing this task and core questions.

Author contributions J.M.F. and A.C. devised and further developed the modelling approach, respectively. J.M.F., A.C. and T.S.I. validated the modelling approach. J.M.F., A.C. and T.S.I. implemented the analysis codes. A.C., J.M.F. and B.Y.H. designed the experimental procedure and downstream analysis. D.L.P., G.P.B., N.G., M. Hasen and S.A.S. performed the intracranial stereotactic surgery. V.K., M. Hegazy, A.M.G. and L.L. provided neurological care for the study participants. J.M.F., A.C. and B.Y.H. drafted the manuscript. J.M.F. and A.C. devised and further developed the neural analyses, respectively. J.M.F., A.C. and T.S.I. implemented them. A.C., T.S.I., M.C.F., E.A.M. and A.G.C. collected the sEEG data. G.D.S., A.C. and E.A.M. collected the behavioural data. E.B., N.R.P. and A.W. contributed to data collection. E.B., N.R.P., A.W. and S.B.M.Y. contributed to data interpretation.

Funding This project was supported by NIH grants R01 DA038615, R01 MH125377, U01 NS121472, and R01 MH129439, and by the McNair foundation.

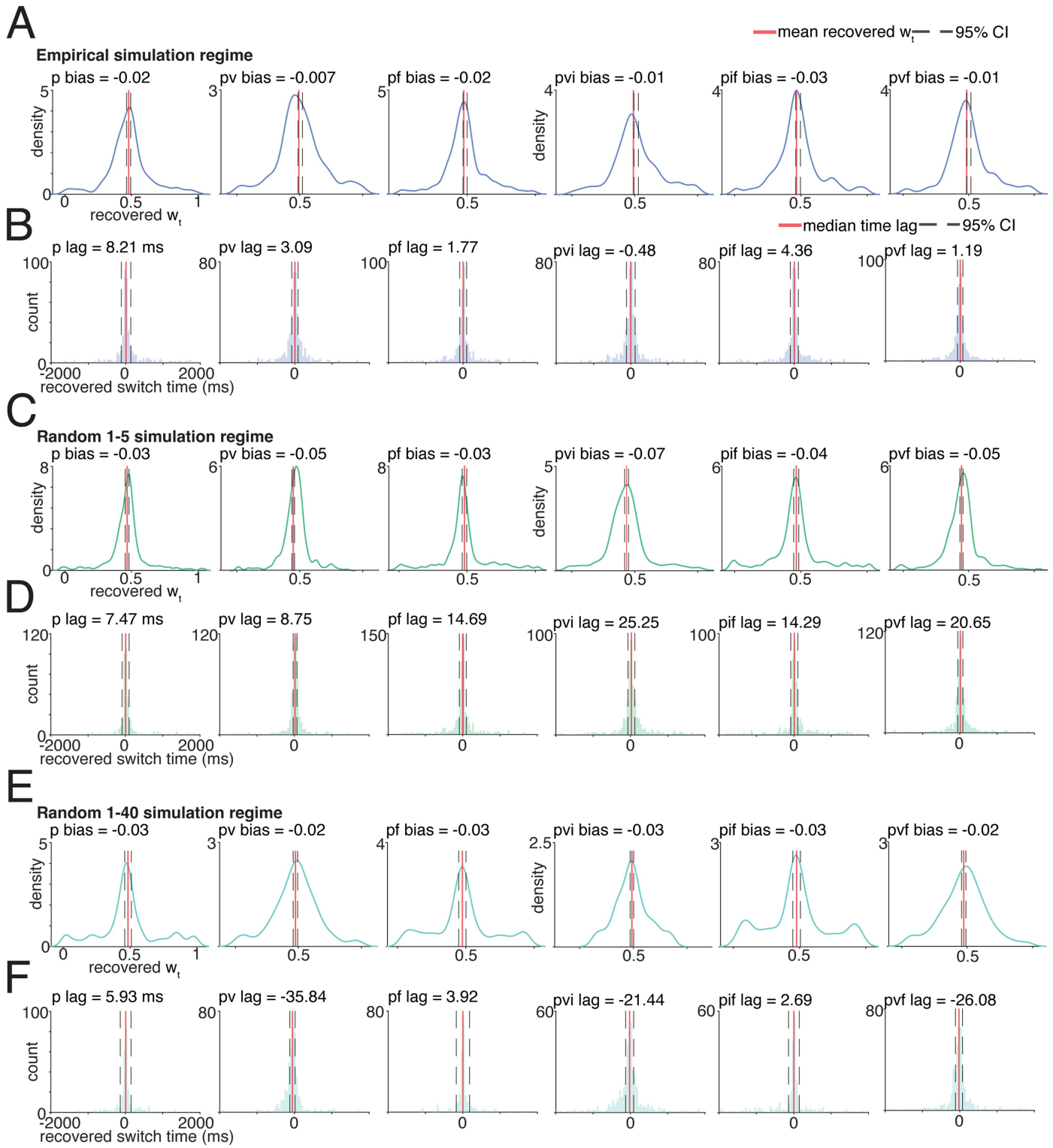
Competing interests S.A.S. is a consultant for Boston Scientific, Abbott, Koh Young, Neuropace, Zimmer Biomet and co-founder of Motif Neurotech. The other authors declare no competing interests.

Additional information

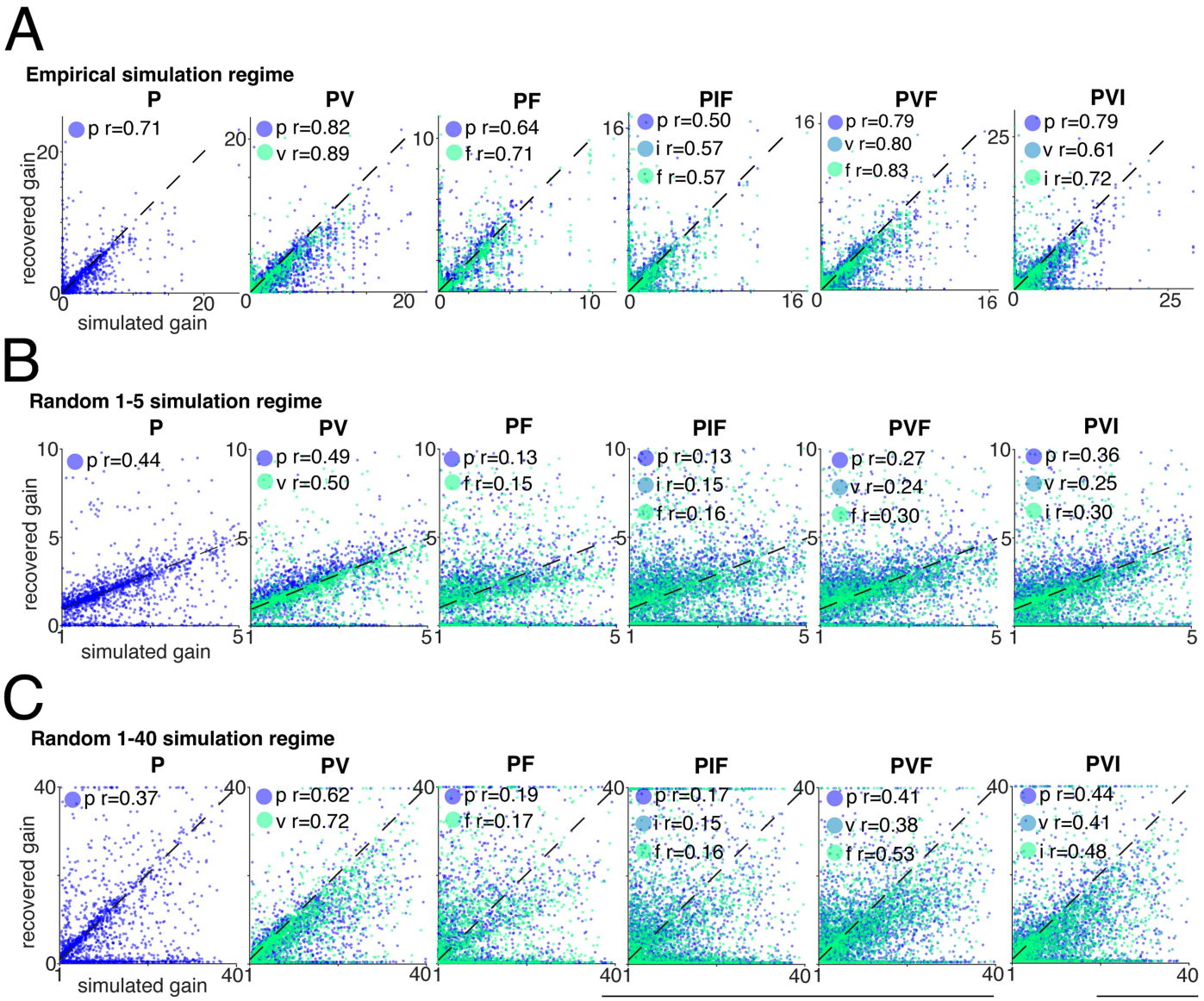
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-026-10896-8.

Correspondence and requests for materials should be addressed to Benjamin Y. Hayden.

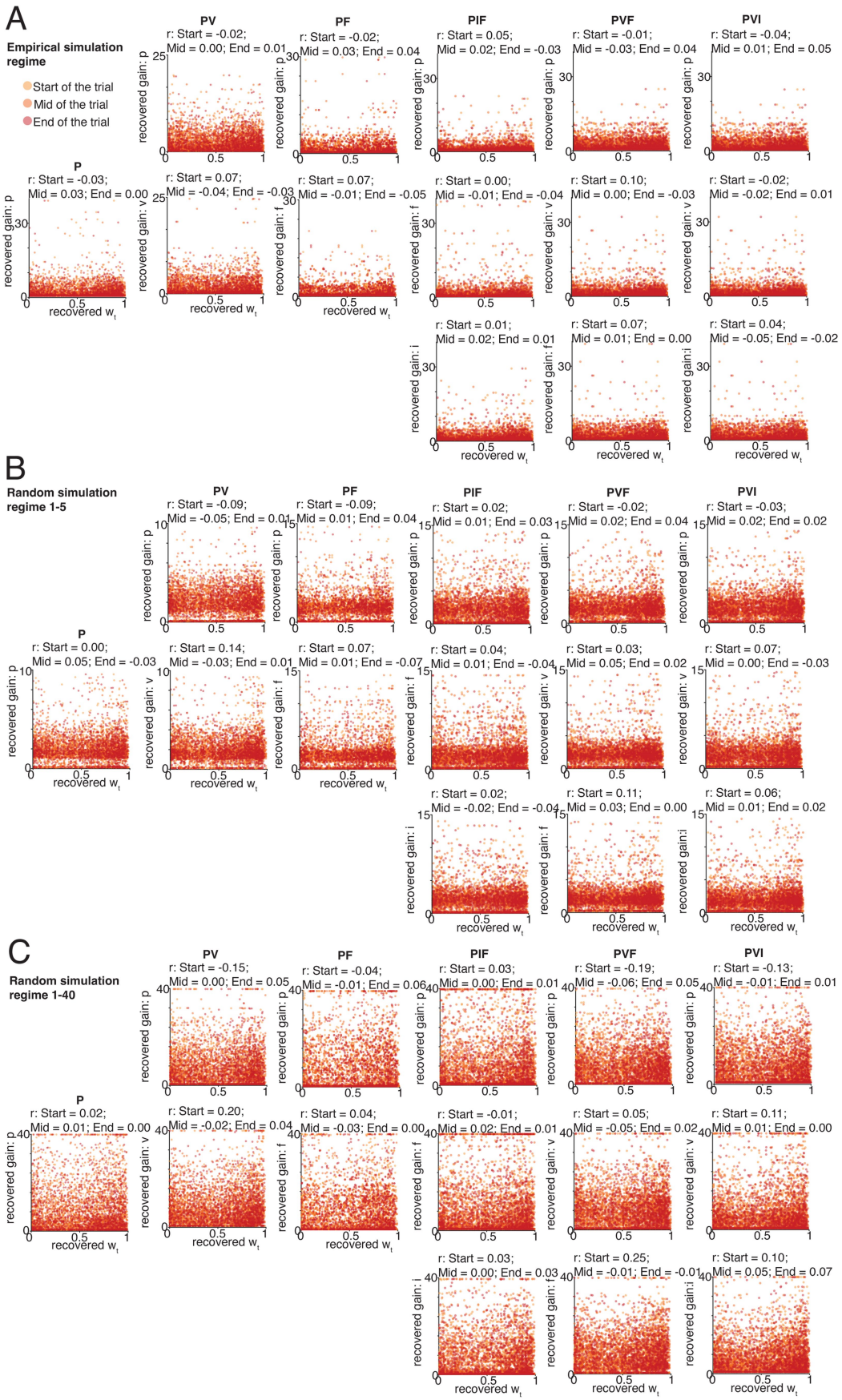
Peer review information *Nature* thanks Sebastian Musslick and the other, anonymous, reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | Model recovery of w_t value and time at switch point ($w_t = 0.5$, $t = 0$ ms). (**A**). Empirical simulation regime (n = 30 simulations × 30 trials × 2 prey-controllers): distribution of recovered w_t values at the true switch point ($w_t = 0.5$), quantifying bias in switch magnitude recovery across controller models. Red line indicates the median recovered value; dashed lines indicate the central 95% interval of the recovered distribution. (**B**). Distribution of temporal lag between recovered and true switch times for $w_t = 0.5$, quantifying timing accuracy of switch detection across controller models. Red line indicates the median lag; dashed lines indicate the 95% CI. (**C**–**D**). Same as (**A**–**B**), but for the random 1-5 simulation regime. (**E**–**F**). Same as (**A**–**B**), but for the random 1-40 simulation regime.

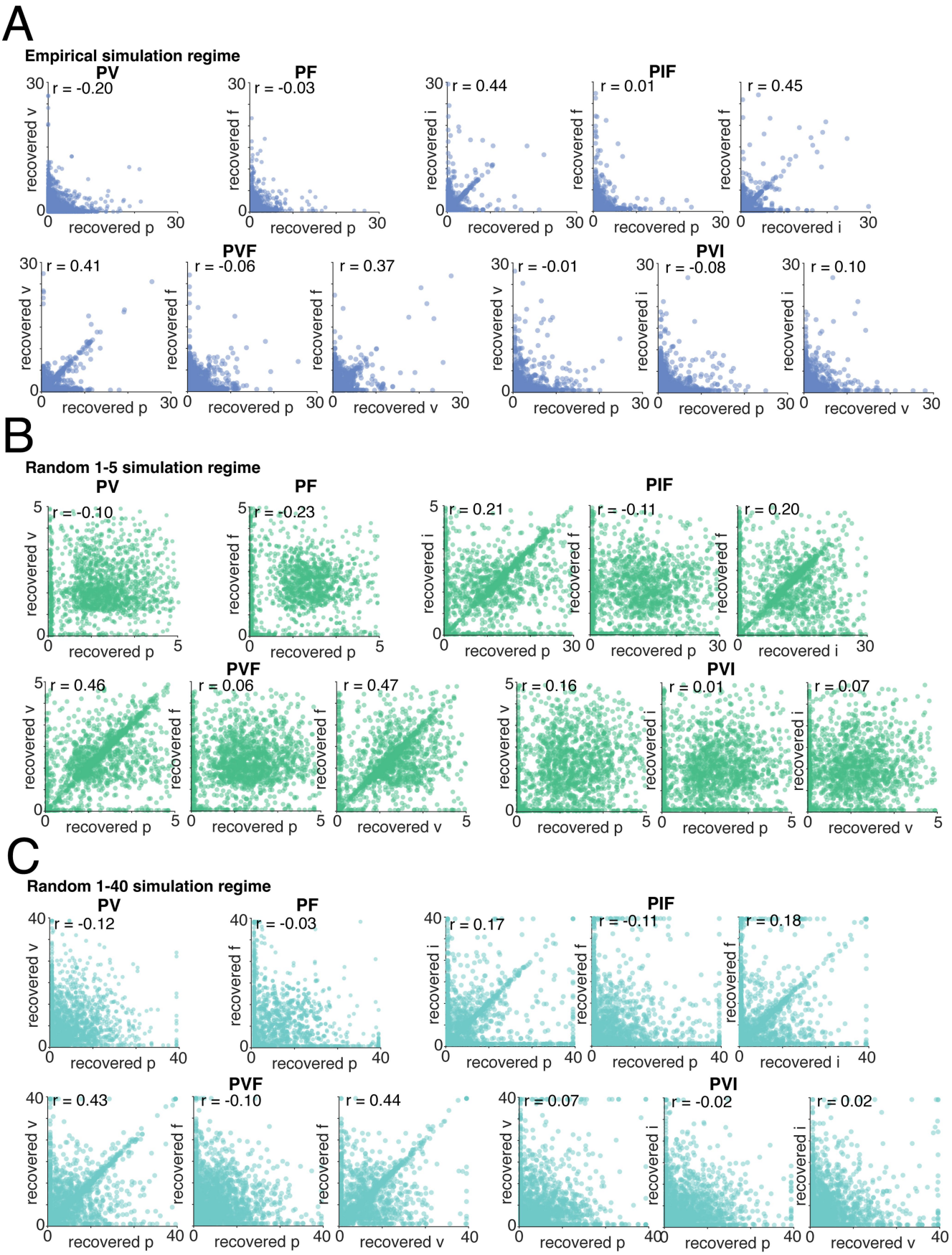


Extended Data Fig. 2 | Model recovery of controller gains across controller classes. (**A**). Empirical simulation regime (n = 30 simulations × 30 trials × 2 prey-controllers): scatter plot showing the relationship between the simulated gains and the gains recovered from the model. Each point represents one recovered gain form a simulated trial. Each column represents a controller class. (**B**). Same as panel A, but for the random 1-5 simulation regime. (**C**). Same as panel A, but for the random 1-40 simulation regime.



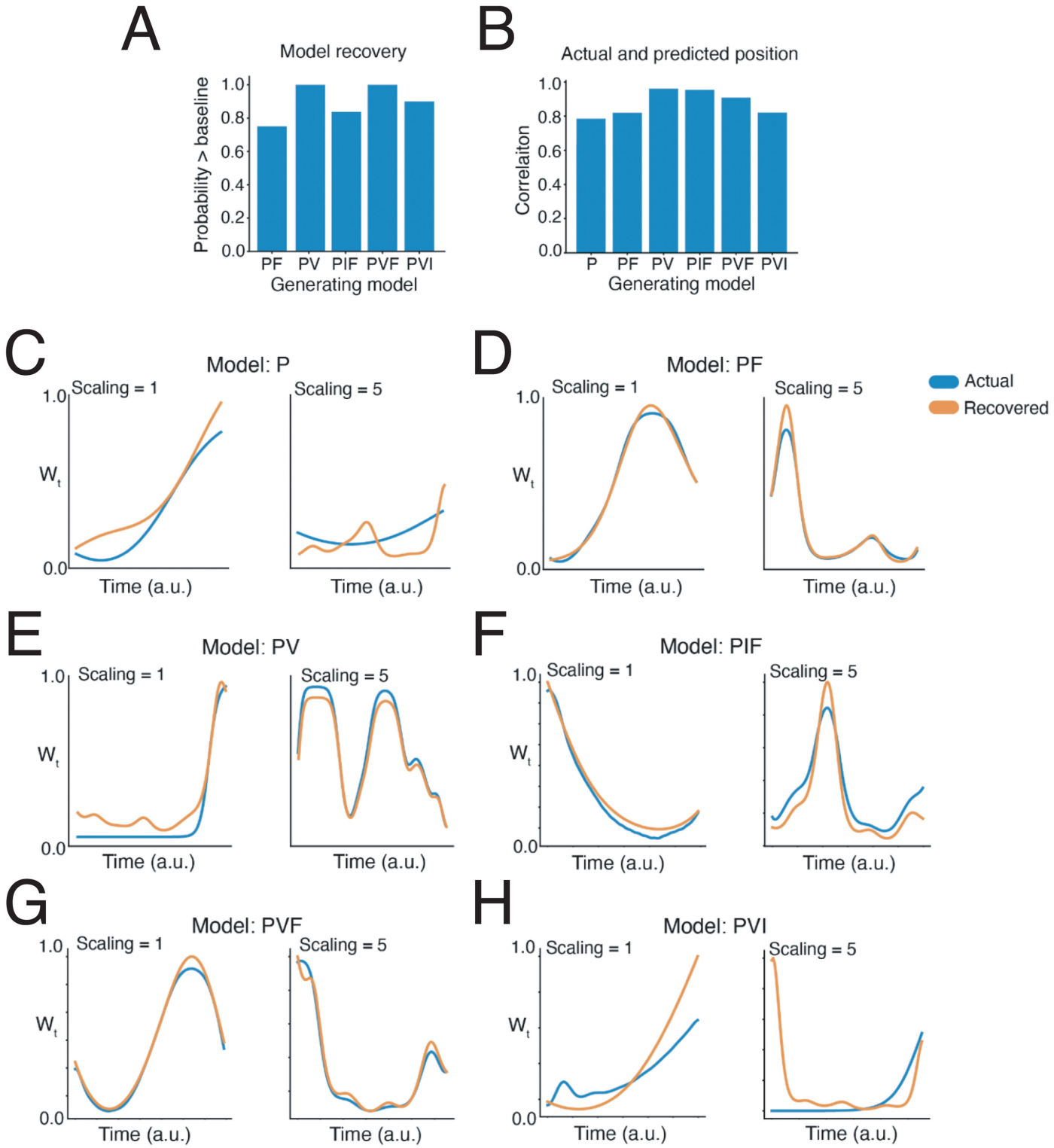
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Relationship between recovered controller gains and recovered w_t trajectories across controller classes. (**A**). Empirical simulation regime (30 simulations × 30 trials × 2 prey-specific controllers): scatter plots showing the relationship between recovered controller gains and recovered $\mathbf{w}_t$ trajectories at the start, middle, and end of the trial. Each point represents one recovered gain estimate from one prey-specific policy in a simulated trial. Each column represents a controller class. Pearson's correlation coefficient (r) is shown for each trial epoch. (**B**). Same as panel A, but for the random 1-5 simulation regime. (**C**). Same as panel A, but for the random 1-40 simulation regime.

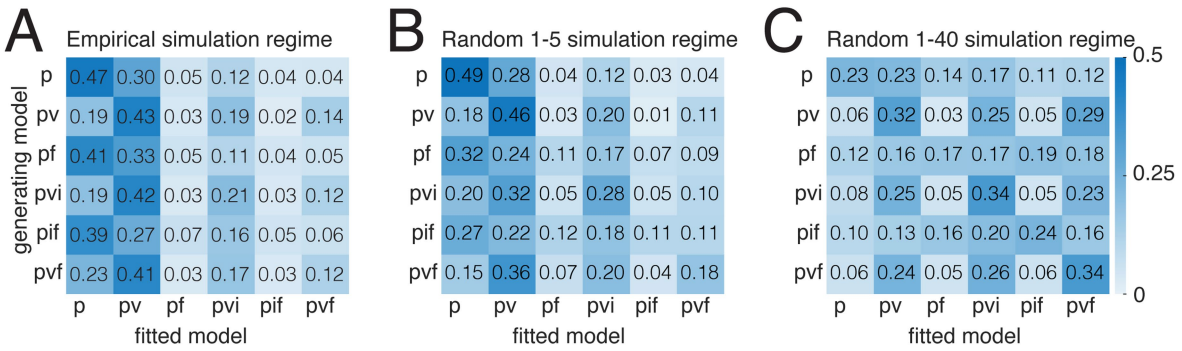


Extended Data Fig. 4 | See next page for caption.

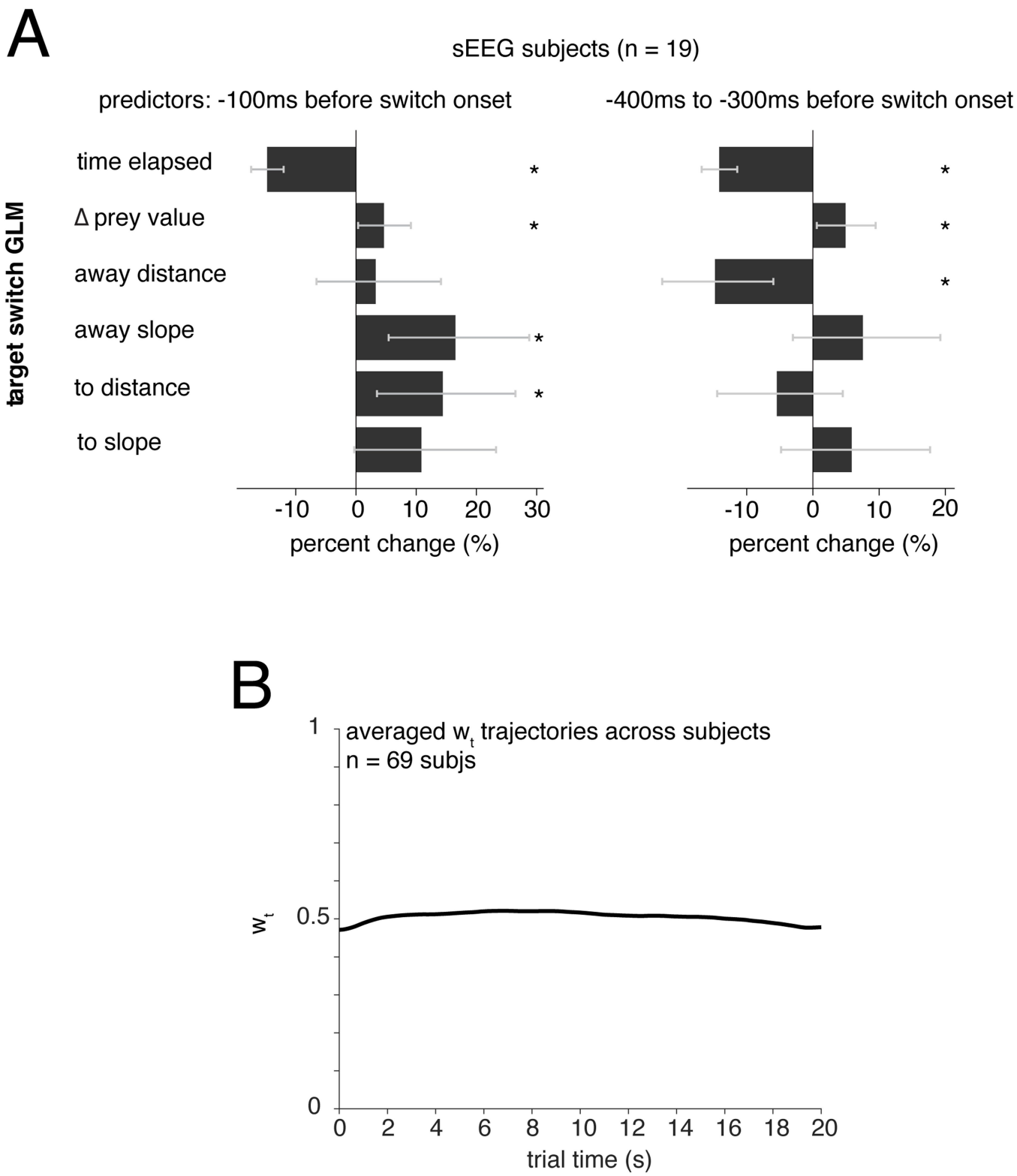
Extended Data Fig. 4 | Pairwise relationships between recovered controller gains across controller classes. (**A**). Empirical simulation regime (30 simulations × 30 trials × 2 prey-specific controllers): scatter plots showing pairwise relationships between recovered controller gains across controller classes. Each point represents one recovered gain estimate from a simulated trial. Pearson's correlation coefficient (r) is shown for each gain pair. (**B**). Same as panel A, but for the random 1-5 simulation regime. (**C**). Same as panel A, but for the random 1-40 simulation regime.



Extended Data Fig. 5 | Model recovery of controller class and blending parameter $\mathbf{w}_t$. We simulated ground-truth trajectories using known controller class (*p*) and known $\mathbf{w}_t$ profiles, considering five types of $\mathbf{w}_t$. Simulations were generated by sampling 20 trials across sessions, with trial durations matched to the empirical distribution (approximately 0.91, 1.45, 2.83, and 4.1 s). To fix model complexity, simulations were fit using a fixed number of basis functions, rather than the model-averaging procedure used in the main text. (**A**). Posterior probability of recovering the true generative controller model relative to a baseline position-only controller. Across all controller classes, the true generative model was recovered above baseline (all probabilities > 0.5). Controller classes are defined in Extended Data Table 3. (**B**). Correlation between true and predicted position trajectories for each generative controller class, averaged across simulated trials, showing good recovery across classes (mean r = 0.83). (**C**–**H**). Example simulations showing recovery of the ground-truth blending parameter $\mathbf{w}_t$ for each controller class. Blue lines indicate true $\mathbf{w}_t$, orange lines indicate recovered $\mathbf{w}_t$. Left panels (scaling = 1) show low-complexity trajectories, and right panels (scaling = 5) show higher-complexity trajectories. Complexity was controlled by the scale parameter of the Gaussian process used to generate $\mathbf{w}_t$. Across models, the mean recovery of $\mathbf{w}_t$ was r = 0.87.



Extended Data Fig. 6 | Full model recovery confusion matrix. To assess overall model identifiability, we performed a full confusion analysis across three simulation regimes. For each generative model corresponding to a controller class (Extended Data Table 3), we simulated datasets with known parameters (30 datasets with 30 trials each), fit all candidate models to each dataset, and selected the best-fitting model using the evidence lower bound (ELBO). (**A**–**C**). Confusion matrices quantifying how reliably each generative controller class could be distinguished from the others under the (**A**). empirical simulation regime, (**B**). random 1-5, and (**C**). random 1-40 simulation regimes. Recoverability was significantly above chance for controller classes ***p***, ***pv*** and ***pvi*** (binomial test, all $p < 0.0001$), whereas the remaining classes showed weaker identifiability. The *pv* model, which provided the best fit to participant behavior, was significantly recovered (43% correct; 95% CI [0.40, 0.47], $p < 0.0001$), but was frequently confounded with ***pvi*** and ***pvf***. This overlap likely reflects the nested structure of the controller classes, indicating that interpretation of fitted model identities should account for partial non-identifiability between related controller families.



Extended Data Fig. 7 | Behavioral predictors of target switching and average controller dynamics. (**A**). Target switch GLM restricted to sEEG subjects (n = 19), showing the contribution of behavioral predictors at two time windows before the switch onset. Left: predictors computed as the mean value within the 100 ms preceding switch onset (−100 to 0 ms). Right: predictors computed as the mean value from an earlier window preceding switch onset (−400 to −300ms). Bars indicate the mean posterior effect size (percent change), and error bars indicate the 95% credible interval, as in Fig. 2h. Asterisks indicate predictors whose 95% credible interval excluded zero. (**C**). Averaged $\mathbf{w}_t$ trajectory across all the participants (n = 69).

Extended Data Table 1 | Experimental cohorts and participants demographics

Cohort	Number of participants	Mean age ± SD	Gender (M/F)	Collection site	Collection period	Task apparatus
sEEG subjects	19	40.1 ± 13.1	8/11	Baylor College of Medicine	Dec 2022 – July 2025	Computer screen + commercially available joystick (Logitech – extreme 3D pro)
Healthy subjects	33	23.5 ± 4.5	19/14	Baylor College of Medicine	Nov 2024 – July 2025	
Healthy subjects	17	22.2 ± 4.5	8/9	University of Minnesota	Jan 2022 – Feb 2023	

Extended Data Table 2 | Number of recorded neurons across participants and brain regions

subject	ACC	hippocampus	OFC	para hippocampus	entorhinal	amygdala	thalamus	PCC
1	3	6	0	0	0	0	0	0
2	15	38	4	0	0	0	0	0
3	18	25	0	0	0	0	0	0
4	25	16	0	0	0	0	0	0
5	25	21	0	0	0	0	0	0
6	0	17	0	0	0	0	0	0
7	0	22	0	0	0	0	0	19
8	7	42	0	0	0	0	0	0
9	12	41	0	0	0	0	0	0
10	17	44	0	0	0	0	0	0
11	25	58	0	0	0	0	0	0
12	5	0	23	0	0	0	0	0
13	9	8	0	34	0	0	0	0
14	0	39	11	17	0	0	8	0
15	31	44	29	0	0	0	0	0
16	24	38	12	0	0	15	12	0
17	20	43	0	0	0	22	0	0
18	12	27	8	0	0	21	13	0
19	6	17	16	0	17	11	9	0
neurons count	254	546	103	51	17	69	42	19

Extended Data Table 3 | Controller types

Equation	Description	Associated Model Name
$u_{pred} = g_p\, e_p$	Current position error	P
$u_{pred} = g_p\, e_p + g_v\, e_v$	Current position and velocity errors	PV
$u_{pred} = g_p\, e_p + g_f\, e_f$	Current and future position errors	PF
$u_{pred} = g_p\, e_p + g_f\, e_f + g_i \int e_p$	Current, future, and integral position errors	PFI
$u_{pred} = g_p\, e_p + g_v\, e_v + g_i \int e_p$	Current, velocity, and integral position errors	PVI
$u_{pred} = g_p\, e_p + g_v\, e_v + g_f\, e_f$	Current, velocity, and future position errors	PVF

Error signals are denoted e_p ($e_p = x_{position,\ subject} - x_{position,\ target+k}$) with subscripts indicating type (e.g., e_p is current position error).

nature portfolio

Corresponding author(s): Benjamin Hayden

Last updated by author(s): Jun 3, 2026

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	Behavioral data were collected using custom task software written in MATLAB (MathWorks; versions R2019b and later) utilizing Psychtoolbox-3. All participants performed the task using a Logitech Extreme 3D Pro joystick, and behavioral data were acquired at 60 Hz. Human intracranial recordings were obtained using Ad-Tech Behnke-Fried stereotactic depth electrodes (8 microwires per probe; Ad-Tech Medical) implanted using a ROSA ONE Brain robotic system (Zimmer Biomet). Neural activity was acquired using a 512-channel NeuroPort System (Blackrock Microsystems) sampled at 30 kHz. Electrode locations were verified using co-registered pre-operative MRI and post-operative CT scans and reconstructed using FreeSurfer (v7.4.1), BioImage Suite (v3.5), iELVis, and RAVE. Single-neuron activity was isolated using Wave_clus spike sorting (Chaure et al., 2018) followed by manual curation based on waveform shape, refractory-period violations, and inter-spike interval distributions.
Data analysis	Data analyses were performed using custom codes written in MATLAB (MathWorks; R2024b) and Python 3.11. Python analyses used NumPy (v2.3.0); SciPy (v1.15.2), scikit-learn (v1.6.1), JAX (v0.6.0), NumPyro (v0.18.0), statsmodels (v0.14.2), Linderman Lab ssm package (v0.0.1). Demixed principal component analysis followed the implementation described by Kobak et al. (2016). All custom analysis codes are publicly available at: https://github.com/BCM-Neurosurgery/PreyPursuit

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

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The complete preprocessed behavioral and neural datasets generated during this study have been deposited in a Figshare repository and are publicly available at: https://doi.org/10.6084/m9.figshare.32572764

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	Biological sex was recorded for all participants and is reported in Table S1 and S2. Sex was determined from medical records for the intracranial cohort, while it was self reported for the healthy behavioral cohort. The intracranial cohort included 19 participants (8 male, 11 female). The healthy behavioral cohort included 50 participants (27 male, 23 female). No analyses were performed as a function of sex because the study was not designed or statistically powered to evaluate sex-related differences.
Reporting on race, ethnicity, or other socially relevant groupings	No socially constructed or socially relevant categorization variables (e.g., race or ethnicity) were used in this study. Participants were characterized only by clinical and demographic covariates (age, sex, epilepsy diagnosis for the intracranial cohort) relevant to the neural recordings.
Population characteristics	Nineteen epilepsy patients undergoing clinical sEEG monitoring participated. Age ranged from 20 to 68 years (mean ± sd = 40.1 ± 13.1 years). All participants were patients with drug-resistant epilepsy implanted with depth electrodes for clinical seizure localization. Periods containing seizures were excluded from all analyses. 50 healthy participants were included for behavioral analysis only. Age ranged from 18 to 33 years for the Baylor cohort (mean ± SD = 23.5 ± 4.5); age ranged from 18 to 34 years for the Minnesota cohort (mean ± SD = 22.2 ± 4.5). Gender for both cohorts is reported above.
Recruitment	Intracranial participants were recruited from patients with drug-resistant epilepsy already undergoing clinically indicated sEEG monitoring at Baylor College of Medicine; recordings were obtained opportunistically during this clinical stay. Participation was voluntary and did not affect clinical care. Because the sample is limited to surgical epilepsy patients, it is not representative of the general population, and electrode placement was determined solely by clinical need rather than research aims, which constrains anatomical coverage. These constraints are inherent to human single-neuron research and are not expected to bias the within-subject effects reported here. Participants received a standard monetary compensation, which did not depend on task performance. All participants received the same compensation. All healthy participants provided informed consent in accordance with protocols approved by the respective institutional review boards at each site. Also here, participants received standard monetary compensation that did not depend on task performance.
Ethics oversight	The study was approved by the Institutional Review Board of Baylor College of Medicine, Houston, TX (protocols H-18112 for the intracranial cohort; protocol H-54759 for healthy subjects). All procedures followed the Declaration of Helsinki. Participants gave written informed consent and were reminded of their right to withdraw at any time.

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	No formal statistical method was used to predetermine sample size. For the intracranial dataset, sample size was determined by the availability of eligible patients undergoing clinically indicated stereotactic EEG monitoring and meeting inclusion criteria. During the review process, six additional participants were included, increasing the final cohort to 19 participants and strengthening the robustness of the reported findings. This sample size is comparable to, or larger than, those commonly reported in human single-neuron recording studies. The healthy behavioral cohort comprised 50 participants and was collected to provide an independent behavioral comparison dataset. The final

sample sizes yielded sufficient numbers of trials, switch events, and recorded neurons to support the statistical analyses reported in the manuscript.

Data exclusions

No participants were excluded from the study. At the neuronal level, exclusion criteria were pre-established and applied during preprocessing. Only well-isolated single units meeting standard spike-sorting criteria were retained, including consistent waveform shape, refractory-period violations below 5%, and appropriate inter-spike interval distributions. All neurons passing these criteria were included in the analyses; analyses were not restricted to task-responsive neurons. Periods containing seizures were excluded from all analyses. For specific analyses, neurons with zero total firing or fewer than two valid observations were excluded because the relevant statistical quantities could not be computed. No healthy participants were excluded.

Replication

The reproducibility of the findings was assessed using multiple complementary approaches. Behavioral results were evaluated in both the intracranial cohort and an independent cohort of 50 healthy participants performing the same task, yielding consistent behavioral patterns. Neural analyses employed cross-validation, held-out test datasets, permutation testing, and repeated resampling procedures where appropriate. During peer review, the intracranial cohort was expanded from 13 to 19 participants; the inclusion of these additional participants did not alter the principal conclusions of the study. The principal conclusions remained unchanged across these validation procedures.

Randomization

Randomization into experimental groups was not relevant to this study. This was an observational within-subject design in which all participants performed the same behavioral paradigm (the prey-pursuit task), and no participants were assigned to different experimental conditions or treatment groups. Consequently, no participant-level randomization procedures were performed.

Blinding

All spike sorting and preprocessing were performed prior to the analyses, ensuring that neurons isolation and inclusion were determined prior to the statistical analyses. Blinding to group allocation was not relevant because the study did not involve experimental groups; all participants performed the same task and all neurons were analyzed using the same predefined processing and analysis pipelines. Statistical inference relied on permutation tests, shuffle controls, and cross-validation procedures rather than subjective experimenter judgment.

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

Plants

Seed stocks

N/A

Novel plant genotypes

N/A

Authentication

N/A