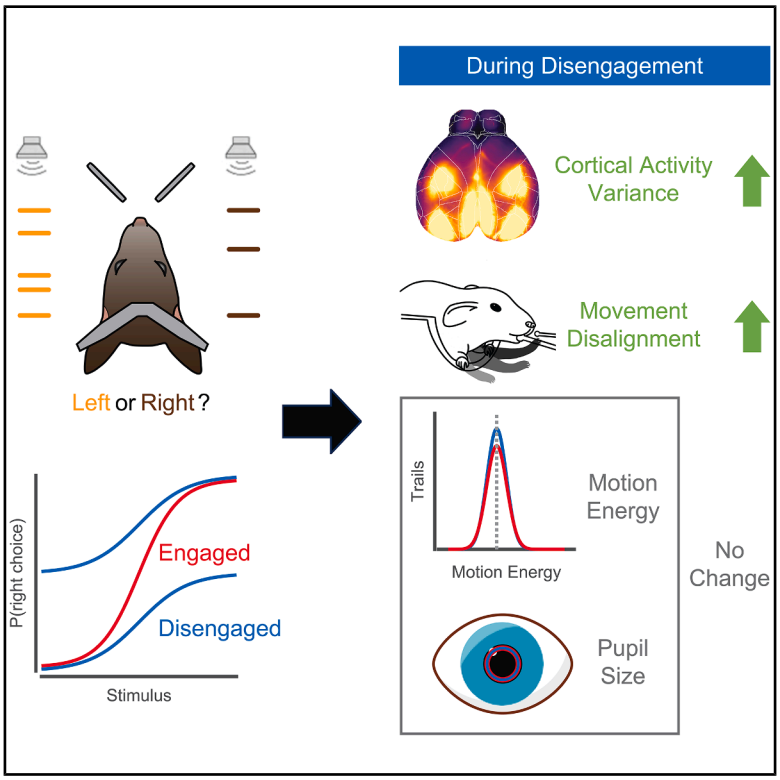
Neuron

Article

Spontaneous movements and their relationship to neural activity fluctuate with latent engagement states

Graphical abstract



Highlights

- Behavioral modeling identifies engaged and disengaged states in mice making decisions
- During disengagement, trial-to-trial variability in cortical activity is elevated
- Enhanced neural activity follows spontaneous movements during disengagement
- Disengagement alters movement stereotypy but not motion energy or pupil diameter

Authors

Chaoqun Yin (尹超群), Maxwell D. Melin, Gabriel Rojas-Bowe, ..., Alex Kostiuk, Simon Musall, Anne K. Churchland

Correspondence

achurchland@mednet.ucla.edu

In brief

Yin and Melin et al. show that disengaged decision-making (defined by biased choices) is coupled with higher neural variance, stronger movement-related cortical activity, and less stereotyped spontaneous movement, indicating a deep connection between behavioral states, cognition, and movement.

Yin et al., 2025, Neuron *113*, 3048–3063
September 17, 2025
https://doi.org/10.1016/j.neuron.2025.06.001

CellPress

CellPress
OPEN ACCESS

Neuron

Article

Spontaneous movements and their relationship to neural activity fluctuate with latent engagement states

Chaoqun Yin (尹超群),[1,2,9] Maxwell D. Melin,[1,2,3,9] Gabriel Rojas-Bowe,[1,2] Xiaonan Richard Sun,[4] João Couto,[2] Steven Gluf,[5] Alex Kostiuk,[1,2,3] Simon Musall,[6,7,8] and Anne K. Churchland[2,10,*]

[1]UCLA Neuroscience Interdepartmental Program, Los Angeles, CA 90095, USA
[2]Department of Neurobiology, University of California, Los Angeles, Los Angeles, CA 90095, USA
[3]UCLA-Caltech Medical Scientist Training Program, Los Angeles, CA 90095, USA
[4]Section of Neurosurgery, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, USA
[5]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
[6]Institute of Biological Information Processing (IBI-3), Forschungszentrum Jülich, 52428 Jülich, Germany
[7]Department of Systems Neurophysiology, Institute for Zoology, RWTH Aachen University, 52074 Aachen, Germany
[8]Faculty of Medicine, Institute of Experimental Epileptology and Cognition Research, University of Bonn, 53127 Bonn, Germany
[9]These authors contributed equally
[10]Lead contact
*Correspondence: achurchland@mednet.ucla.edu
https://doi.org/10.1016/j.neuron.2025.06.001

SUMMARY

Switching between cognitive states is a natural tendency, even for trained experts. To test how cognitive states impact neural activity and behavior, we measured cortex-wide neural activity during decision-making in mice. During disengagement, neural activity was more variable across trials and could be better explained by a linear encoding model. This increase in explained variance during disengagement was associated with two changes: modestly stronger neural encoding of movements generally and an increase in task-independent movements specifically. Surprisingly, behavioral videos showed similar motion energy in both cognitive states. But while the overall amount of movements remained similar, movement alignment changed: as animals slipped into disengagement, their movements became less stereotyped. These idiosyncratic movements were a strong predictor of task performance and engagement. Taken together, our results suggest that the temporal structure of movement patterns constitutes an embodied signature of the cognitive state with a profound relationship to neural activity.

INTRODUCTION

Defining how neural circuits interpret sensations to guide decisions is a key step in understanding brain function.[1,2] Decisions with uncertain sensory information are especially intriguing because they rely on sophisticated computations such as accumulating evidence, combining sensory information with prior knowledge, and ultimately making inferences about the state of the world.[3–5] Understanding the neural mechanisms underlying these computations has been challenging, in part because the variability of cortical neurons makes single-trial measurements hard to interpret.[6,7]

Two approaches have been deployed to cope with the variability of neural responses. The first approach aims to minimize sources of variability by limiting the movements that experimental subjects are permitted to make. Examples of this strategy include training non-human primates to fixate on a visual target and training rodents to hold their heads in the same position during the stimulus presentation period.[8–11] Although this approach reduces the variability of neural activity, it leads to decision-making behaviors that differ critically from most natural situations in which spontaneous movements are frequent.[12–14] Therefore, even basic questions, such as whether spontaneous movements are associated with good or bad decision-making performance, remain unanswered.

The second approach to cope with the variability of neural responses is to collect neural activity during many repeats of the same decision and then to average these. This approach relies on the assumption that the relationship between neural activity and behavior remains stationary over an experimental session. However, the relationship between neural activity and behavior is often not stationary but instead state-dependent. For instance, the activity of agouti-related peptide (AgRP)-expressing neurons in the hypothalamus depends critically on an animal's hunger state,[15] unselectively averaging AgRP activity collected during hunger and satiated states would obscure the state-dependent activity of these neurons. The ability to perform large-scale

recordings partially circumvents this problem because analyses can focus on single-trial activity from simultaneously recorded neurons (and thus all from a given state).[16–18] However, it is not always clear when heterogeneous population responses are due to innate circuit properties or instead are due to random noise and variability across the population. Thus, the interpretation of such population activity and a consideration of how it evolves trial-to-trial still benefit from understanding state-dependent changes in the relationship between neural activity and behavior.

For cognition, the assumption of a stationary relationship between neural activity and behavior has recently been challenged. Behavioral modeling confirms what casual observation of experimental data collection has suggested: animals fluctuate in and out of engagement states over the course of a single experimental session.[19–22] However, the impact of these fluctuating states on neural activity has only just begun to be studied, and the cortex-wide impact of these states remains, like the impact of movements on decision-making, an unanswered question.

We discovered an unexpected link between these two unanswered questions. We measured neural activity across the dorsal cortex while mice made perceptual decisions about uncertain stimuli. We deployed a behavioral model to assign each decision in the experimental session to either an engaged or disengaged state[19] and discovered changes in neural activity across states. Disengagement is characterized by an increase in the trial-to-trial variability of neural activity, while the cortical activity during disengagement can be better explained by spontaneous movements. These changes come with two effects on both the movement-neural activity relationship and the movement pattern itself. First, the impact of spontaneous movements on neural activity, though idiosyncratic, was overall higher in the disengaged state. Second, a closer examination of behavioral videos surprisingly uncovered that as animals became disengaged, the rate and magnitude of movements remained similar, but the movements became less stereotyped and temporally uncoupled to task events. We defined a metric of this temporal alignment to task events, task-independent movement (TIM), and found it to be a strong predictor of task performance, even more so than pupil diameter, an extensively used arousal indicator.[23–25] Taken together, these results argue that movements define cognitive states that are distinct from arousal and explain state-dependent changes in cortical activity, indicating a deep connection between movement, behavioral state, and cognition.

RESULTS

Mice occupy engaged and disengaged latent states during auditory perceptual decisions

We trained four mice to perform an auditory discrimination task. Mice were head-fixed to a behavioral apparatus with speakers on the left and right sides of the animal's head (Figure 1A, left). Mice initiated trials by grabbing a set of handles in front of them. The 0.5 s long window before the trial initiation defined the baseline epoch. Poisson-distributed clicks were then played through the left and right speakers at randomized times.[11,12,26] After a variable delay epoch (0–0.5 s), two spouts moved in, and the mouse reported its decision by licking, and licks to the side with the higher click rate were rewarded (Figure 1A, right; STAR Methods, task 1). The mean rate of left vs. right clicks was systematically varied to alter the difficulty of trials. EMX-Cre-GCaMP6s was expressed in all pyramidal neurons, allowing us to perform simultaneous wide-field imaging[26] (STAR Methods). Once mice were trained on all stimulus difficulties, we computed the psychometric function for individual mice (Figure 1B, light gray lines) and for all mice (Figure 1B, black line). Psychometric function parameters for individual subjects are shown in Figure S1.

To determine the latent behavioral states of mice, we fit a hidden Markov model with Bernoulli generalized linear model observations (GLM-HMM) to the behavioral data (Figure 1C). This class of model extends HMMs, which are widely applied to time-varying data to extract latent dynamics.[19,22,27–29] The GLM-HMM uses stimulus sensitivity and response bias to assign trials to discrete states that last tens to hundreds of trials. We combined data from the four mice shown in Figure 1B to ensure sufficient trial counts for training the model (STAR Methods). Optimal training parameters, including the number of states, were found via grid search with 10-fold cross-validation. Test log-likelihood continuously increased with the addition of more latent states and reached a saturation at $\sim$0.25 bits/trial (Figure S2A).

We chose a 3-state model because additional states provided only marginal improvements to test log-likelihood, similar to previous studies.[19,22] The 3-state model identified one engaged state with high stimulus weight and two disengaged states (right and left biased, Figure 1D). After fitting the GLM-HMM, we used the forward-backward algorithm to compute the probability that a mouse was in a particular state for a given trial. We then assigned trials to three groups, identified by their most likely state, and computed the psychometric function for each group (Figure 1E; STAR Methods). As expected, the engaged state showed the lowest bias and the highest slope. Disengaged states were marked by noticeable rightward or leftward bias and a smaller slope. Fluctuating state probabilities for an example session are shown in Figure 1F, demonstrating state transitions that span tens to hundreds of trials. The total fractional occupancies were highest for engaged and right-biased states, with fewer trials being assigned to the left-biased state (Figure S2B). For subsequent analyses, we grouped trials from the 3-state model into either engaged or disengaged states (where trials from left- and right-biased states were combined). This ensured that any state-dependent changes in neural activity that we observed were not simply due to the animal favoring one choice over the other.

The mean, but not the variance, of cortical activity is similar across states

We next tested whether these latent states might modulate neural activity. We imaged calcium signals through the intact, cleared skull. Fluorescence data were spatially aligned to the Allen Mouse Common Coordinate Framework v3 (CCF).[30] Data were temporally aligned to task epochs (baseline, handle-grab, stimulus, delay, and response) to compute individual trial responses and trial-averaged responses.

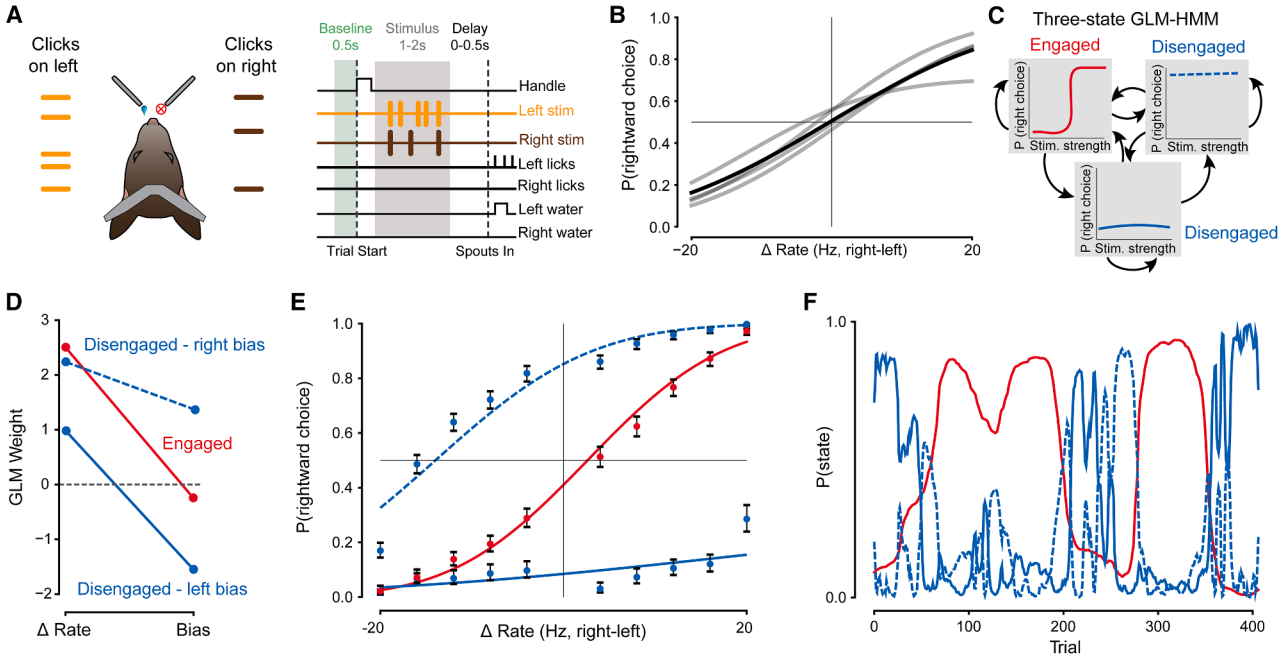


Figure 1. A GLM-HMM identifies discrete behavioral states with distinct psychometric functions

(A) Head-fixed mice perform a spatial auditory discrimination task. Mice are instructed to lick to the high-rate side after a short delay period.

(B) Psychometric curves for 4 mice (gray lines) and average psychometric curve across all mice (black line).

(C) Example schematic of GLM-HMM. Each latent state has a distinct psychometric curve. Black arrows denote transitions between states.

(D) Parameters from the GLM-HMM fit to the four mice shown in (B). GLM weights for each of the 3 latent states are shown, revealing one engaged and two biased states.

(E) Psychometric curves for each latent state identified by the GLM-HMM. Engaged state: bias = 0.131, slope = 1.73. Disengaged states: bias = 0.936 and −0.699, slope = 0.617 and 1.49. Error bars indicate Wilson binomial confidence intervals. Same color conventions as in (D).

(F) Recovered state probabilities from an example session, revealing fluctuations in engagement throughout a session. Same color conventions as in (D).

To understand the impact of engagement on neural activity, we first examined the neural responses of trials assigned to the engaged or disengaged state (Figure 2A). In the somatosensory cortex, upper limb area (Figure 2B), we found that in both states, the trial-averaged neural activity fluctuated as mice initiated trials, formed decisions, and obtained rewards (Figure 2B, thick traces). As in previous work, these modulations likely reflect the frequent, uninstructed movements that the animals make during each of these task epochs. Considerable single-trial variability was also apparent (Figure 2B, thin traces). For this and subsequent analyses, we matched the ratio of correct to incorrect outcomes for the trials that were studied in each state (Figure 2C). This ensured that any state-dependent differences we observed in neural activity were not due to a mismatch in reward-related activity between the two states. Trials in the disengaged state were often either right- or left-biased, and pooling these ensured that both left- and right-choice trials were included in the analysis.

We then compared mean neural activity across states for the entire dorsal cortex (Figure 2D). We considered four epochs in the trial: baseline, stimulus, delay, and response (see gray boxes in Figure 2B). The mean activity for engaged (Figure 2D, top) and disengaged (Figure 2D, middle) trials was similar at most timepoints, with only transient differences between the two (Figure 2D, bottom). A closer look at the temporal dynamics of mean activity in somatosensory and visual cortices showed that trial-averaged neural activity on engaged trials was slightly elevated during the delay epoch (Figure 2E, red trace above blue trace during "delay").

Having observed only modest state-dependent differences in the trial-averaged response (Figures 2D and 2E), we then computed the across-trial variance. This analysis is important because features that modulate neural activity at different moments in each trial will be obscured when many trials are averaged together.[31,32] Interestingly, the cross-trial variance on disengaged trials was reliably higher in some areas (Figure 2F). This means that single-trial activity is more variable in the disengaged state than in the engaged state, a feature that is evident in single-trial responses (Figure 2B, compare thin blue/red traces). This state-dependent difference was evident across multiple task epochs. Differences were most pronounced in the primary motor cortex (MOp) and somatosensory cortex (the upper limb area, lower limb area, and barrel cortex, Figure 2G) and were primarily due to increased variance along the first latent dimension of neural activity (Figure S3).

Our counterbalancing correct and incorrect trials ensures that these findings are not due to reward rate differences (Figure 2C). These findings were also not explained by state-dependent left/right-choice imbalances: when equalizing left and right-choice trials within each state, the engaged state was again associated

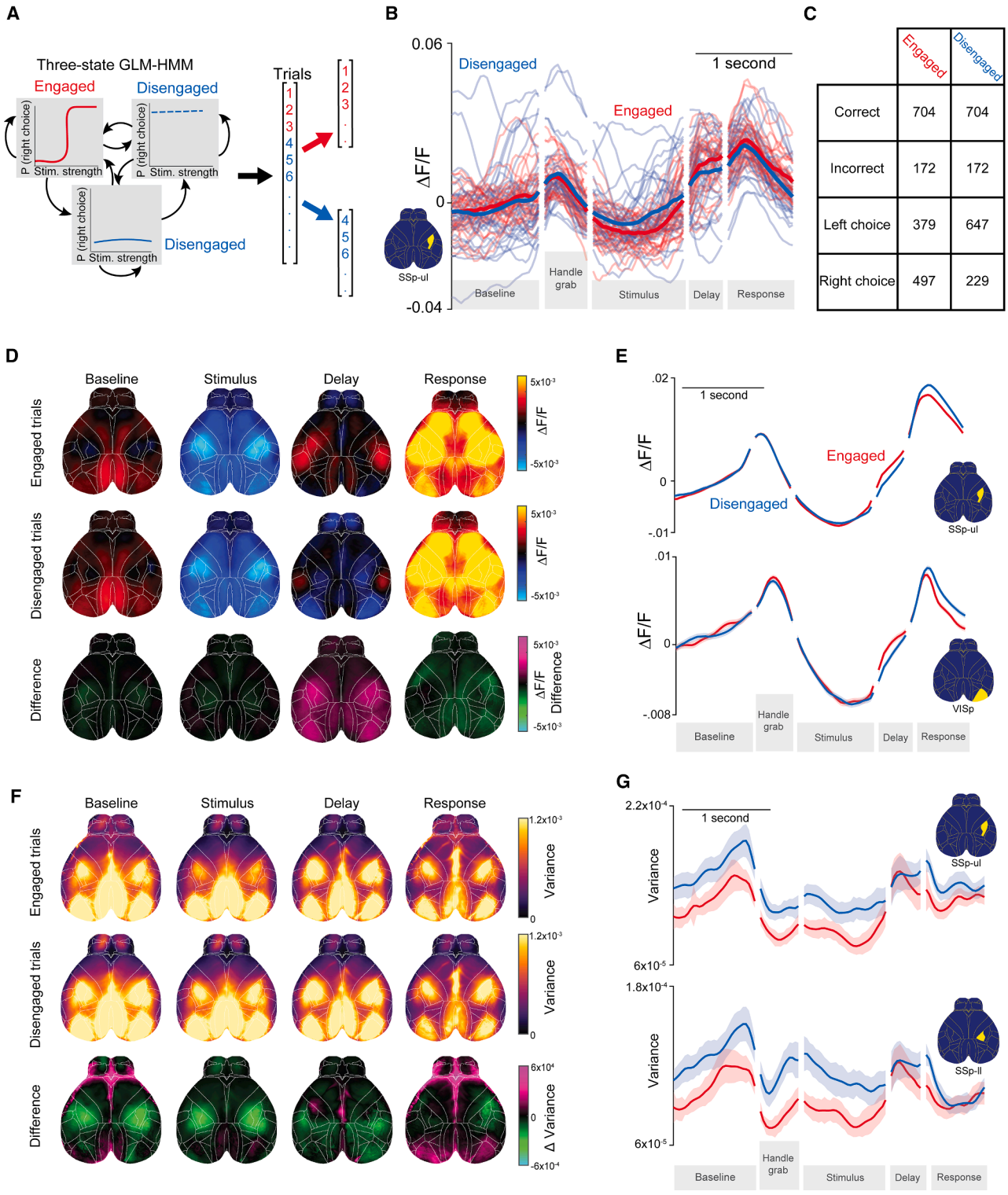


Figure 2. Trial-average responses measured with wide-field imaging are similar, but single-trial variability is altered during engagement vs. disengagement

(A) Estimated states from the GLM-HMM were used to group trials into engaged and disengaged states.

(B) Calcium activity from somatosensory cortex-upper limb area (SSp-ul) aligned to task events for an example session. Red lines denote engaged trials, and blue lines denote disengaged trials. Thick lines: mean trial-averaged activity. Thin lines: individual trials.

(legend continued on next page)

with less trial-to-trial variability (Figure S4). On a smaller subset of sessions that contained a high number of left-choice trials, we computed across-trial variance for each state using only left-choice trials. Again, we observed increased trial-to-trial variance in the disengaged state, indicating this effect is not due to pooling left-biased and right-biased states (Figures S5A and S5B). Taken together, these results show that cortical activity is less variable during the engaged, high-performance state, independent of the state-related task performance differences.

Encoding of movements, but not task variables, is altered during disengagement

To uncover what factors contributed to increased variability during disengagement, we built a linear encoding model.[12] This model was designed to reconstruct neural activity with many regressors, allowing opportunities to uncover differences between the engaged and disengaged states. These regressors included task variables, discrete movements, and video data (see Table S1 for a detailed breakdown of regressor groups). We trained separate models on trials that the GLM-HMM assigned to each state (Figure 3A). Engaged- and disengaged-state models were matched within sessions to have equal amounts of training data. Using these models, we could reliably reconstruct single-trial activity across the dorsal cortex (Figure 3B). To assess overall model performance in each state, we determined the match between the model's prediction and the actual neural data, using data held out from training (cross-validated R^2, Figure 3C). Surprisingly, cvR^2 was higher for models trained on disengaged trials (Figure 3C, left, blue line above red line; Figure 3D, two right columns), meaning that the model could account for more single-trial variance in neural activity when animals were disengaged.

To assess the unique contribution of state, we then assessed the performance of a modified encoding model trained on all trials that included behavioral state as a regressor. Model performance was indistinguishable from a model without a state regressor (Figure 3D, two left columns; see also Figure S5C). This is unlike previous studies using more compact models in which the inclusion of behavioral state improved model fits dramatically compared with models with only sensory inputs.[33] Here, the similar performance on models with vs. without a state regressor demonstrates that state-dependent differences (e.g., Figure 3C) must be driven by other regressors, such as task variables or movements.

We set out to determine which regressors drove the observed state-dependent differences (Figure 3C). Simplified, task-variable-only models performed similarly for the two states (Figure 3E, left; note that the modest values for unique explained variance of task variables are in keeping with previous results[12,26]). By contrast, uninstructed-movement-only models performed differently for the two states (Figure 3E, right, blue line above red). This indicates that uninstructed movements (i.e., movements not required for task performance, e.g., whisking) contribute to the state-dependent differences observed in the full model (Figure 3C). Observing state-dependent differences for the uninstructed-movement-only but not the task-variable-only model (Figure 3E) also argues against the possibility that motion-induced imaging artifacts were larger during disengagement. If motion artifacts were larger during disengagement, the explained variance of task and movement variables would both have changed.

Previous work demonstrated that movements can contribute to both single-trial and trial-averaged activity because the space of all movements includes task-aligned and task-independent components (Figure 3F, compare top vs. bottom rows).[12] We therefore wondered if the observed state-dependent differences were driven primarily by task-aligned movements (i.e., neural variance that can be explained by task variables *or* movements due to correlation between the two) or movements independent of task (i.e., neural variance that can only be explained by movements, which are uncorrelated to task variables). Existing methods[12] can distinguish the variance accounted for by these two types of movements: when movements and task variables are correlated, they account for overlapping cortical variance in linear models (STAR Methods). This correlation can be exploited by comparing the performance of task-variable-only models with task-variable-plus-movement models. The additional variance explained when including movements represents task-independent variance (variance due to movements that was not accounted for in task-variable-only models). Because movement-driven variance must be either task-aligned or task-independent, the task-aligned variance is computed by subtracting task-independent variance from the total variance explained by movements.

Task-aligned movements explained small, equal amounts of cortical variance in engaged and disengaged states (Figure 3G, top row; compare red/blue dashed traces; see STAR Methods). However, when we compared model performance on the movements that could not be accounted for by task variables, we found a consistent difference between the two states (Figure 3G, bottom row, blue line above red line).

We extended our analyses to another cohort of five mice with pyramidal tract (PT) neurons genetically labeled by Fezf2-CreER (key resources table).[26] PT neurons are a subset of EMX-expressing pyramidal neurons. They project to subcortical structures, such as the pons and thalamus.[26,34] Similar to EMX-expressing neurons described above, the variance and cvR^2 of PT neurons were higher during disengagement (Figure S7). The

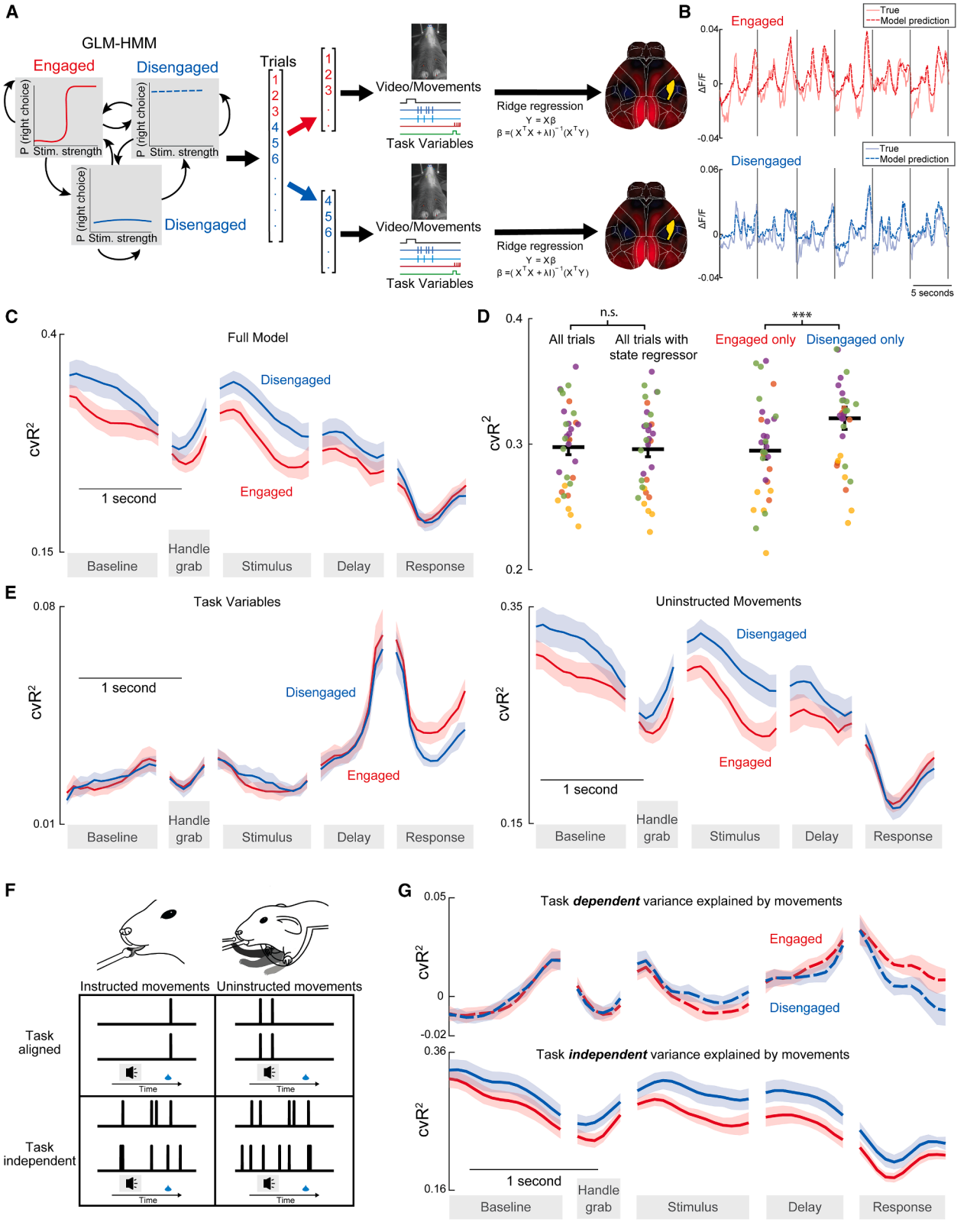
(C) Trial numbers for the analyses shown in (D)–(G). Rewarded/unrewarded trials are counterbalanced to ensure state-dependent differences are not driven by reward.

(D) Heatmaps of average activity during each trial epoch, averaged over 19 sessions from 4 mice. Top row: engaged trials. Middle row: disengaged trials. Bottom row: difference (engaged activity minus disengaged activity). Purple indicates more activity during the engaged state, and green indicates more activity during the disengaged state.

(E) Average activity over time for two example regions. Error: ±SEM across sessions.

(F) Heatmaps of average variance across trials. Higher values denote more trial-to-trial differences in neural activity. Row conventions are the same as in (E).

(G) Across-trial variance plotted for two example regions. Error: ±SEM across sessions.



(legend on next page)

cell bodies of PT neurons are primarily located in layer 5b, indicating that the results reported above remain consistent in the deep cortical layers.[34]

This analysis suggests that the impact of engagement on neural activity is related to the movements independent of the task. Two mechanisms could underlie this effect: (1) the same movement is associated with a stronger neural response during disengagement, or (2) the movement pattern itself changes across engaged and disengaged states. We tested these two possibilities in the following analyses.

The relationship between movements and cortical activity changes modestly with disengagement

Our encoding models, trained separately on engaged and disengaged states, generated time-varying beta weights for each regressor (a.k.a., the kernels) within each state. These beta weights specify the relationship between cortical activity and movements. We examined the kernels for movement regressors by computing the L1-norm of the kernel over time.

The kernels of a subset of event-based regressors were significantly (but modestly) larger during disengagement (Figure S6). We also tested the kernels of behavior video data, and the kernels of some singular value decomposition (SVD) components differed across states (Figure S6D). To summarize these differences, we calculated the L1-norm of the 200 component kernels and again found that the kernels during disengagement were modestly larger (Figure S6E). Overall, our findings suggest that some movements are associated with slightly stronger cortical activity during disengagement, though the changes are idiosyncratic and modest at the individual regressor level.

Overall movement magnitude is consistent across states

To better understand movements in the engaged vs. disengaged states, we developed a new analysis based on the movement of individual body parts. We employed DeepLabCut, a neural-network-based movement tracking software,[35] to track the movements of 27 body parts (Figure 4A). We used DeepLabCut instead of dimensionality-reduced videos or raw pixel values because these latter methods are difficult to relate to specific animal movements.[36] The movement trajectories of DeepLabCut-tracked body parts varied over trials and time, though velocities are unchanged across states (Figure 4B, two example body parts).

To ensure that DeepLabCut tracking captured similar movement information to our video-based approach, we again trained task-variable-only and uninstructed-movement-only models, this time using DeepLabCut tracking data as regressors (Figure S8). As before, uninstructed movement explained more neural variance during disengagement, while the neural variance associated with task variables remained the same across both states. This test reinforces the robustness of our findings and confirms that DeepLabCut tracking captures key movement information.

To test whether mice move more during disengagement, we quantified the motion energy (cumulative position change across frames) of all labeled body parts in the stimulus and delay epochs. To prevent the total motion energy from being dominated by body parts that naturally have longer moving distances, such as the forelimbs, we standardized (Z score) the motion energy of different body parts separately before averaging them. In most trials, standardized motion energy tended to increase gradually from the beginning of the trial to the moment when the water spouts came in and the mouse reported its choice (Figure 4C). Interestingly, motion energy was very similar during engagement and disengagement (Figure 4D, left). The probability of being in the engaged state (P(engaged)) and motion energy were uncorrelated (Figure 4D, right). This argues that the increase in explained variance during disengagement (Figures 3C and 3E, right) did not arise because disengaged animals move more.

Stereotyped movement patterns are disrupted during disengagement

Although motion energy is unchanged across states, we noticed that the trajectories of labeled body parts were more stereotyped trial-to-trial during engagement compared with disengagement (Figure 4B, top; red traces are more similar to each other than blue traces). This could occur if, for instance, movements became desynchronized to task events during disengagement. Indeed, our analysis of neural data hinted that the movements independent of task change during disengagement (Figure 3G).

Figure 3. More single-trial variance can be explained in the disengaged state, primarily because of variance due to task-independent, uninstructed movements

(A) Schematic describing the modeling approach. Trials were identified as engaged or disengaged with a GLM-HMM. Linear models used video, movements, and task variables to predict neural activity for engaged and disengaged states. Cortical heatmaps on the right denote instantaneous ΔF/F to be predicted by the model. Red indicates positive ΔF/F, and blue indicates negative ΔF/F. Sessions with fewer than 50 trials per state were excluded from analysis.

(B) Example traces from SSp-ul (yellow highlighted region from A) and model predictions for engaged and disengaged trials from one session. Gray vertical lines denote trial starts.

(C) Total variance explained across-trial epochs for engaged and disengaged states. Variance explained was computed for the "full model," including task variables, instructed movements, and uninstructed movements.

(D) Left two columns, total variance explained by linear encoding models trained on all trials, with and without a state regressor. Including a state regressor in the model does not increase variance explained. Error bars denote mean ± SEM. Scatter points denote individual sessions, with colors denoting mice. Linear mixed-effects model with mouse identity as a random effect, $p = 0.81$. Right two columns: linear encoding models trained on engaged or disengaged trials from matched sessions. Linear mixed-effects model with mouse identity as a random effect, $p < 0.001$.

(E) Left, variance explained for a model containing only task variables. Right, variance explained for a model containing only uninstructed movements. Data were combined across sessions with mean ± SEM plotted.

(F) Diagram demonstrating how movements can be viewed as instructed vs. uninstructed or task-aligned vs. task-independent.

(G) Same as (C) and (E), but variance explained is shown for task-dependent (top) vs. task-independent (bottom) movements.

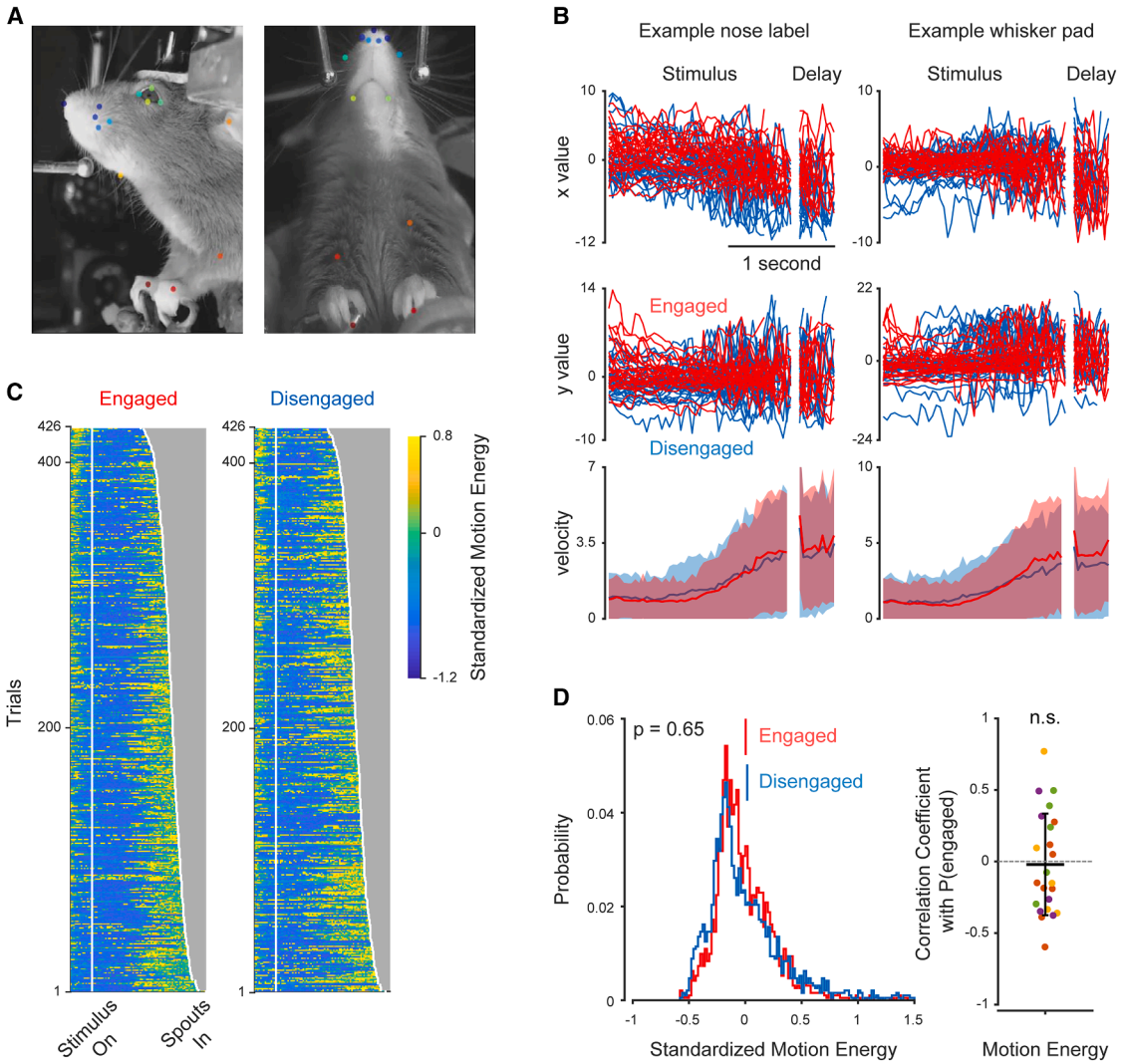


Figure 4. Animals move the same amount in the engaged and disengaged states

(A) The movements of 27 body parts were tracked with DeepLabCut. Footage from two camera angles was used. Colored dots: points tracked by DeepLabCut.

(B) The standardized x, y coordinates and averaged velocity at each video frame of two example DeepLabCut labels (50 randomly selected engaged and disengaged trials). The shade in the velocity plot denotes the standard deviation.

(C) The motion energy of DeepLabCut labels for an example animal in engaged and disengaged trials. Each row is one trial. Trials are sorted based on when the water spouts moved in, allowing the mouse to report a choice ("spouts in"). The trials with the top 20% P(engaged) are designated as "engaged trials," while the bottom 20% of trials are selected as "disengaged trials."

(D) The motion energy is similar across engaged and disengaged states. Left: the distribution of motion energy (averaged over the stimulus and delay epochs for each trial). The red/blue bars denote the mean motion energy of engaged/disengaged states, respectively. Standardized motion energy of engaged trials = 0.0033 (mean) ± 0.3062 (standard deviation), standardized motion energy of disengaged trials = 0.0166 ± 0.3953. 3,836 trials from 4 animals were included. A linear mixed-effects model was fitted to the motion energy with engagement state as a fixed effect and mouse identity as a random effect. No significant difference was detected (p = 0.6456). Right: the correlation coefficient between motion energy and P(engaged) was calculated separately for individual sessions. The 4 colors represent the 4 mice from which the sessions were recorded (same color conventions as in Figure 3D). The horizontal bar and error bar denote the mean coefficient (−0.0206) and standard deviation (±0.3553) across sessions. A linear mixed-effects model was fitted to the correlation coefficients with mouse identity as a random effect. The correlation coefficients are not significantly different from zero (p = 0.7788).

We previously observed that in trained experts (but not novices), animals make well-timed whisks following stimulus onset and reward delivery.[12] To quantify the extent to which movements are aligned in time to task events, we built a model that used task variables (e.g., stimulus strength, upcoming choice) to predict the trajectories of the DeepLabCut labels (STAR Methods). Some movement patterns depend on task variables: as an illustration, in advance of the choice, trial-averaged forearm movements might differ between trials that end in left vs. right choices (Figure 5A, thick gray traces). Other movements are independent

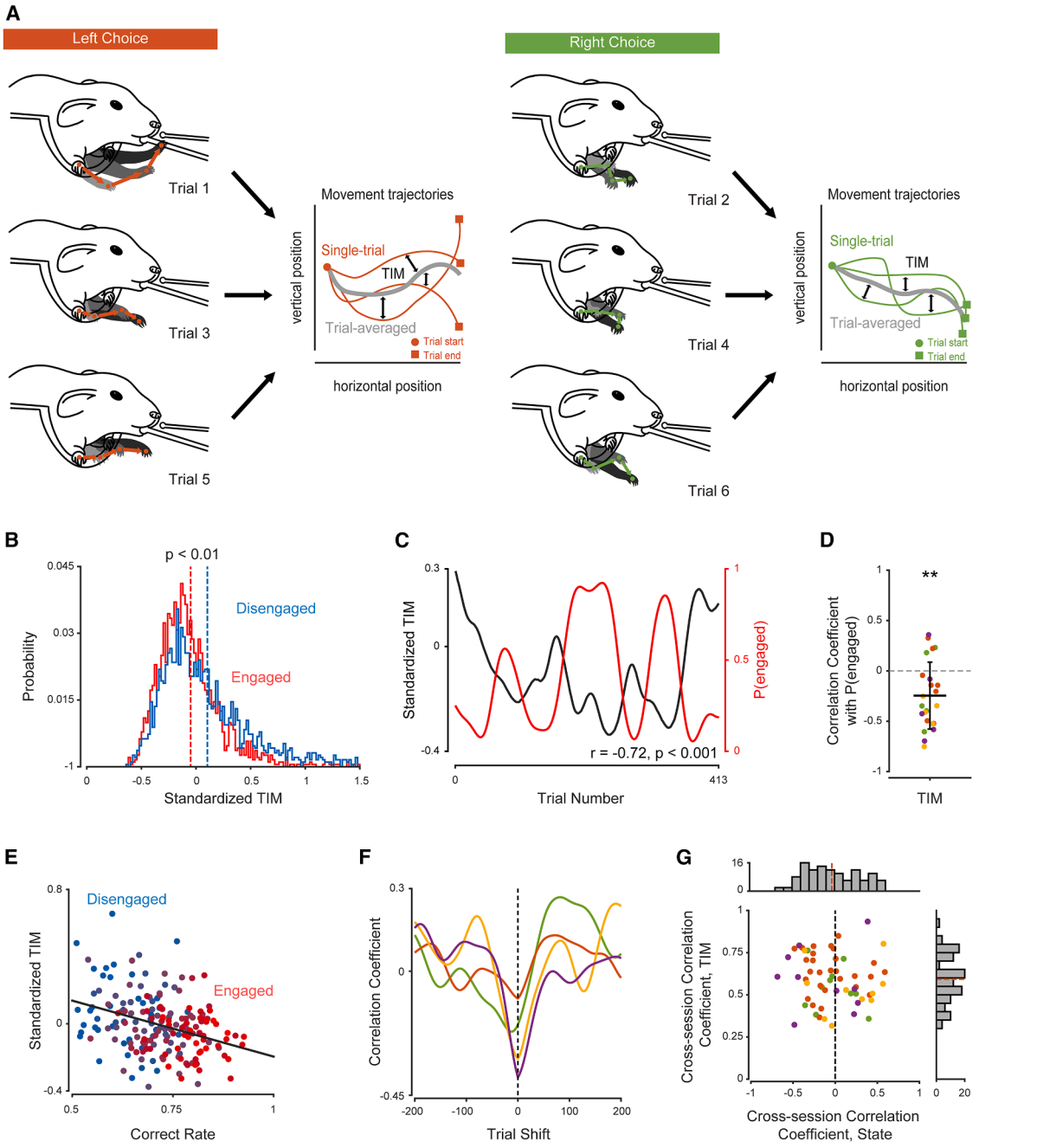


Figure 5. A novel behavioral metric demonstrates a strong relationship between the alignment of movements and the animal's latent state

(A) Schematic example: movements of the hand differ for left and right choices. TIM measures the mismatch between model-predicted stereotyped trajectories (thick traces) and actual single-trial movements (thin traces). TIM is low when spontaneous movements are more stereotyped trial-to-trial.

(B) Distributions of trial-averaged TIM values (stimulus and delay epochs) in the engaged/disengaged states. The red/blue dashed lines: means of the distributions. Standardized TIM of engaged trials = −0.0538 (mean) ± 0.3293 (standard deviation), standardized TIM of disengaged trials = 0.1072 ± 0.4638, 3,836 trials (4 animals). A linear mixed-effects model was fitted to TIM with engagement state as a fixed effect and mouse identity as a random effect. TIM is significantly higher in the disengaged state (p = 0.0045).

(C) Smoothed TIM and P(engaged) are negatively correlated (r = −0.7163, p = 3.3286e−66, one example session).

(D) Correlation coefficients of TIM and P(engaged) were calculated separately for individual sessions after TIM and P(engaged) were smoothed (50-trial-long Gaussian filter). Colors represent individual mice (same color conventions as Figure 4D). Horizontal bar and error bar denote the mean coefficient (−0.2508) and

(legend continued on next page)

of task events. For instance, a hand flexion that is unique to each trial, regardless of choice (Figure 5A, thin orange and green traces). We refer to the difference between the model-predicted trajectories and the actual trajectories as "task-independent movements" (TIMs). TIMs were averaged across all labeled body parts over the stimulus and delay epochs to obtain one value for each trial, capturing the extent to which, on that trial, movements were uncoupled from stereotyped trajectories dependent on task events (STAR Methods; see also Figures 9A).

We then tested whether TIM depended on the engagement states of animals. TIM was overall higher during disengaged trials (Figure 5B). Moreover, when we compared fluctuations in TIM with P(engaged), we found them to be inversely correlated (Figures 5C and 5D). Periods of low TIM, when movements tend to be better aligned to task events, are associated with engagement. Accordingly, the spatial distribution of body labels occupies a larger area in high-TIM trials (Figure S9C). This indicates that body label positions are more variable across trials during disengagement. Given that engagement states were modeled on the animals' choices, we also tested the relationship between task performance, TIM, and P(engaged). All these metrics are correlated (Figure 5E). The engaged state is characterized by both high task performance and low TIM (Figure 5E, red points at the right). We further confirmed the relationship of TIM and task performance with another nine animals that are not included in the GLM-HMM. These mice were used in optogenetic inactivation experiments, which could interrupt continuous HMM state, or had cells labeled other than EMX-expression neurons. The negative correlation is significant and robust (Figure S10). Moreover, we computed the cross-correlation between TIM and animal state by shifting TIM over time (sessions from the same animal were merged). The negative correlations of all four animals peaked around 0, implying that there is a near-zero time lag between TIM and engagement (Figure 5F).

We performed three additional analyses to ensure that the negative correlation between TIM and P(engaged) did not occur by chance.[37] One concern is that the fluctuations of TIM and P(engaged) within each session might have fixed patterns. For instance, TIM might always be higher at the beginning and the end of each session, and P(engaged) might have an opposite pattern. If so, the correlation between TIM and P(engaged) might not reflect a meaningful association. To address this, we calculated the cross-session correlations of TIM and P(engaged) respectively after adjusting all sessions to the same length with 1-dimensional interpolation (e.g., if we had three sessions from one mouse, we calculated the TIM correlation of all three possible session pairs: 1–2, 1–3, and 2–3). If TIM fluctuations followed a fixed pattern, the correlation coefficients of most pairs would be high. The same analysis was applied to P(engaged). Only TIM showed high cross-session correlation (averaged r = 0.59). The cross-session correlation of P(engaged) was minimal (averaged $r = -0.03$) (Figure 5G). This result argues against the possibility that TIM-state correlation emerged from the fixed patterns of both TIM and P(engaged).

A second concern is that, because the linear regression models used to predict the body part positions included outcome-related regressors, the negative TIM-state correlation merely reflects the fluctuation of model-fitting quality with specific outcome-related regressors (e.g., rewarded vs. unrewarded). To address this concern, we recalculated TIM after excluding all outcome-related regressors and repeated the correlation analyses. The correlation structure between task performance, TIM, and P(engaged) was preserved (Figures S11A–S11C).

A third concern is that the increase in TIM may not indicate the animal's engagement in the current trial but instead reflect that the animal experienced an omitted reward and time-out punishment in the previous trial. In this case, TIM would be higher in the disengaged state simply because the disengaged state had more incorrect previous trials. However, we observed that TIM was higher for disengaged than for engaged trials even when the outcome of the previous trial was matched (Figure S11D).

TIM is more tightly linked to engagement state than pupil diameter

Pupil diameter has previously been used to estimate the arousal of subjects across species. Previous studies have reported positive or inverted-U-shaped correlations between arousal level and pupil diameter.[20,25,38–40] We tested whether TIM and pupil diameter contain overlapping state-related information or whether TIM reveals distinct aspects of behavioral states. We extracted the mean pupil diameter in the baseline and stimulus epochs from videos (Figure 6A), then smoothed and downsampled the data with a 50-trial window. The pupil diameter in neither the baseline nor the stimulus epochs correlated with task performance or P(engaged) (Figure 6B). We further tested the correlation between pupil dilation (the pupil diameter difference between the stimulus and the baseline epochs), task performance, and P(engaged). No significant correlations were found either (Figure S12). Our results argue that the states defined by TIM are distinct from the states defined by pupil diameter.

standard deviation (±0.3411) across sessions. A linear mixed-effects model was fitted to the correlation coefficients with mouse identity as a random effect. Correlation coefficients are significantly below zero (p = 0.0015).

(E) Scatter plot shows the correct rate and TIM calculated in an independent 50-trial window. The color grade indicates the P(engaged) value of trials. TIM is negatively correlated with both engagement and task performance (TIM-correct rate correlation: $r = -0.3315$, p = 2.6445e−6; TIM-P(engaged) correlation: $r = -0.2799$, p = 8.4443e−5).

(F) Trial shift test shows that there is a near-zero time lag in the negative TIM-engagement correlation. Each curve indicates, for a single mouse, the correlation between TIM and P(engaged) when they were offset in time (same color conventions as in D). The negative correlation is at its strongest level when the offset is near zero.

(G) Points: the correlation coefficient of a pair of sessions from the same mouse (same color conventions as in D). Marginal histograms: distribution of cross-session correlation of TIM (right) or P(engaged) (top). Red dashed lines: mean of each distribution. The cross-session correlation of TIM is non-zero (0.5988 ± 0.1431, linear mixed-effects model p = 5.8325e−45). The cross-session correlation of P(engaged) is not significantly different from 0 (−0.0316 ± 0.3208, linear mixed-effects model p = 0.2886).

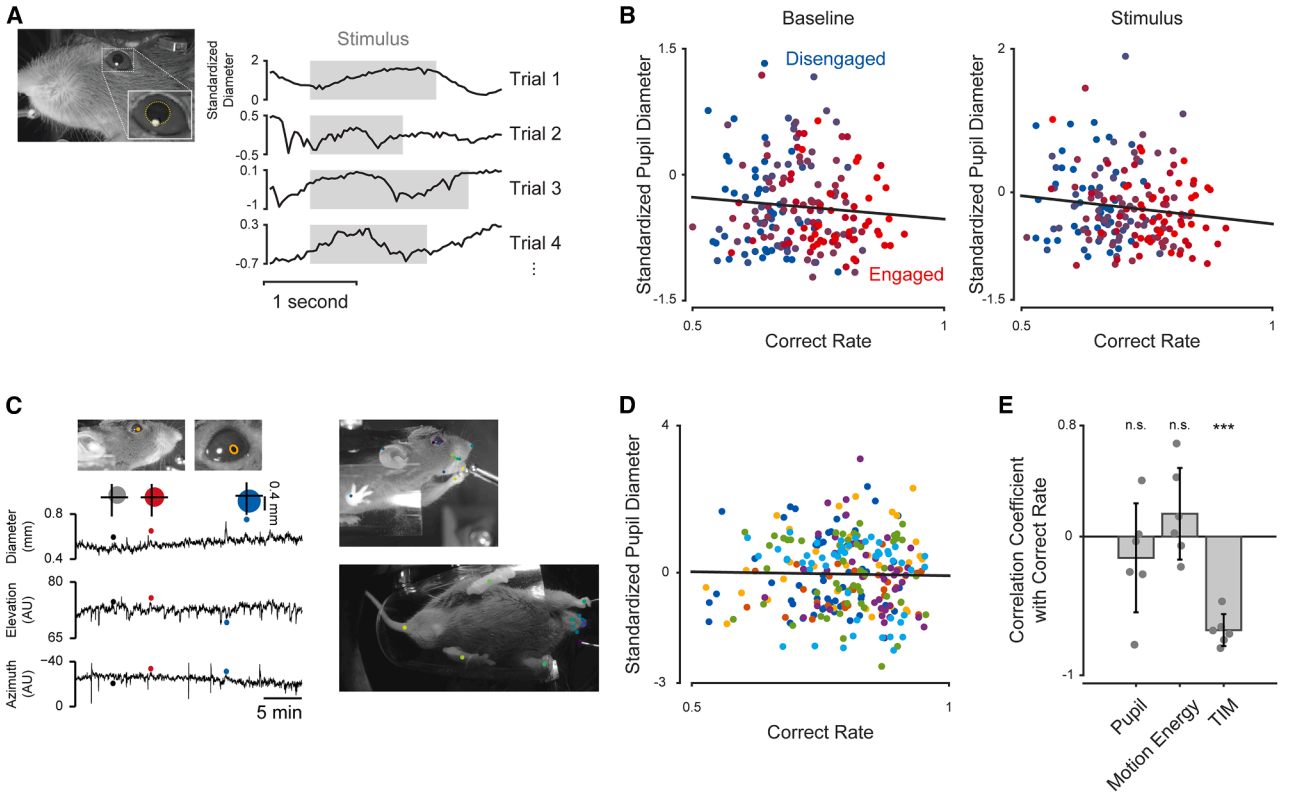


Figure 6. Latent states are correlated to TIM but not pupil diameter

(A) Pupil diameter was measured based on the lateral videos with customized code.[26] The standardized pupil diameters of 4 example trials are shown.

(B) No correlation exists between pupil diameter, task performance, and P(engaged) (baseline epoch: $r = -0.0881$, $p = 0.2257$; stimulus epoch: $r = -0.1325$, $p = 0.0663$). The data from 23 sessions of 4 animals were used. Each point denotes one trial. The color grade indicates the P(engaged) value of trials. Pupil diameter and correct rate were smoothed with a 50-trial-long Gaussian filter and downsampled every 50 trials.

(C) Top left: the pupil diameter was extracted from the lateral videos of head-fixed mice performing a visual task. Pupil diameter was estimated using customized software.[41] Bottom left: filtered pupil diameter (area shown as circles) and eye position (crosshairs represent the center of the eye). Right: 30 body parts in the new dataset were tracked with DeepLabCut for motion energy and TIM calculation. The colored dots show the DeepLabCut labels on two example frames from the videos.

(D) Relationship between correct rate and standardized pupil diameter for 6 mice (colors) that comprise the dataset described in (C). Correct rate and pupil diameter were smoothed with a 50-trial-long Gaussian filter and downsampled every 50 trials to generate the points shown. Pupil diameter and correct rate were not significantly correlated ($r = -0.0505$, $p = 0.3862$).

(E) The correlation coefficients of task performance and pupil diameter/standardized motion energy/standardized TIM are plotted for each animal. Bars indicate the averaged correlation coefficients across animals. Only TIM shows significant correlation with task performance. Pupil-correct rate correlation coefficients = −0.1534 (mean) ± 0.3926 (standard deviation), one-sample t test, $p = 0.3825$; motion energy-correct rate correlation coefficients = 0.1640 ± 0.3301, one-sample t test, $p = 0.2781$; TIM-correct rate correlation coefficients = −0.6743 ± 0.1147, one-sample t test, $p = 2.9174e{-}5$. The data from 6 animals was used.

To further test our findings, we examined a second dataset collected in the context of a different visual decision-making task (STAR Methods, task 2). Again, we tracked the movement trajectories of 30 body parts with DeepLabCut and extracted the pupil diameter in the 0.5 s long window before visual stimuli were displayed (Figure 6C). We avoided using the stimulus epoch in this dataset because pupil diameter may fluctuate independently of task engagement when visual stimuli were presented. We again observed no correlation between the pupil diameter and task performance (Figure 6D). We further tested the TIM-performance correlation and found that, as before (Figure 5), TIM was inversely correlated with task performance (Figure 6E).

The significant relationship between TIM and task performance suggests that movements might be an alternative to pupil diameter to identify internal states that are linked to task engagement. To test the generality of these observations, we applied TIM-performance analysis to data from freely moving rats trained to distinguish the rate of a 1.0 s long multisensory stimulus sequence (Figure S13A; STAR Methods, task 3). We calculated TIM and motion energy of 19 DeepLabCut-labeled body parts (Figure S13B). Again, we observed a significant negative correlation between TIM and task performance (Figures S13C–S13E), confirming that our findings are not limited to head-fixed mice, whose movements during disengagement might be influenced by restraint.

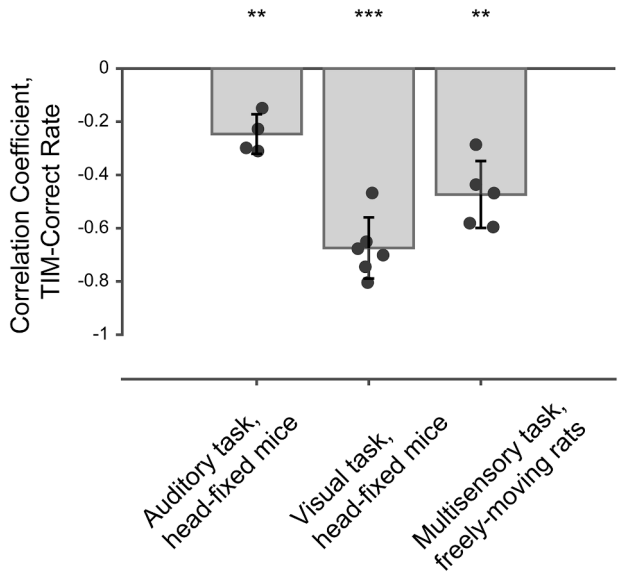


Figure 7. The correlation coefficients between TIM and task performance are plotted for three experimental conditions

Each point denotes the coefficient of one animal. Bars indicate the averaged correlation coefficient across animals. TIM and correct rate are negatively correlated in all 3 datasets that span different task paradigms and species. Correlation coefficients of 4 mice performing the auditory task (task 1) = −0.2463 (mean) ± 0.0743 (standard deviation), one-sample t test, $p = 0.0070$; correlation coefficients of 6 mice performing the visual task (task 2) = −0.6743 ± 0.1147, $p = 2.9174e{-}5$; correlation coefficients of 5 rats performing the multisensory task (task 3) = −0.4734 ± 0.1254, $p = 0.0011$.

Combining the results from three datasets, we consistently observed that TIM is a stable indicator of an animal's performance on cognitive tasks (Figure 7). The consistency of this effect argues that spontaneous movements can be used to infer the animal's engagement across species and contexts.

DISCUSSION

Previous studies have shown that discrete, latent states can govern decision-making dynamics.[19,22] However, the neural and behavioral correlates underlying these states remain poorly understood. In this study, we observed that trial-to-trial neural activity had heightened variability during disengagement. An encoding model accounted for more neural variability during disengagement due to the increased explanatory power of the movement regressors independent of the task (Figure 3). This effect was associated with two neural activity and behavioral changes. First, some uninstructed movements had a modestly larger impact on neural activity during disengagement (Figure S6). Second, although animals moved frequently in both states (Figure 4D), many movements became less stereotyped during disengagement (Figures 5B–5D). Our findings demonstrate that changes in engagement are associated with changes in the movement-neural activity relationship and the temporal structure of spontaneous movements that impact neural activity. They also suggest that movement stereotypy can be used as a powerful, versatile indicator of latent states in diverse experimental conditions, pointing to a surprisingly strong link between movements and cognition.

Existing work has shown that spontaneous movements can have a widespread impact on neural activity during cognitively demanding tasks[12,42,43] and that movements and cognitive functions can be modulated by shared neural circuits.[44] For example, dopaminergic projections from mouse ventral tegmental area and substantia nigra control both movement sequences and cognitive functions (e.g., reward-related learning[45–47]). Correlations between movement and learning were also observed at the behavioral level.[48,49] These findings suggest connections between movement, neural encoding, and cognition but leave the nature of these connections unclear.

Evidence that movements can enhance neural coding comes from locomotion, which is associated with more precise stimulus representations in mouse visual cortex.[50–52] Our work relates to these studies: the decreased neural variability during engagement that we observed could lead to more precise stimulus representations. However, we also found that non-stereotyped, spontaneous movements are associated with disengagement and low task performance. Our results suggest that the relationship between movements and information encoding depends on the nature of the movements. Temporally aligned movements, including rhythmic locomotion, may enhance sensory encoding (or at least leave it undisturbed), but temporally unaligned movements may do the opposite. In our encoding model analysis, only TIMs are associated with stronger neural responses during disengagement, further suggesting the relationships between movement, cognition, and neural encoding can be movement-specific. One potential mechanism is that temporally unaligned movements hinder performance by allowing crosstalk between movement and task-related neural subspaces, driving a confounding overlap between signals.[53–56]

Previous work has uncovered "slow drift" signals that shift the baseline activity, thereby affecting task performance in multiple ways,[33] such as altering the impulsivity level in visual detection.[57] However, our analyses show that a linearly additive "state signal" shared by a large neural population need not be the driving factor underlying engagement state fluctuations (Figure 3D). Instead, engagement state may be largely reflected in trial-to-trial neural variability.

The change in neural variability we observed is reminiscent of prior work reporting state-dependent changes in trial-to-trial variability on multiple timescales: during the seconds before a learned movement,[58] during the minutes while attention is directed to a spatial location,[59,60] and during the weeks-long transition from novice to expert.[61] It is possible that the drop in trial-to-trial variability in those studies is also movement-related. However, a change in movements is likely not the sole explanation because, at least for primate early visual areas, movement modulation of neural activity is weak.[8,62,63] Therefore, the variability drop likely arises, at least partially, from another source. One hypothesis is that neural activity fluctuations, linked with bespoke weights to individual neurons, change with internal state.[64,65] An appealing idea is that these shared fluctuations are present in many species and areas, but the link between

the fluctuations and movements is species/area-specific. Interestingly, recent efforts examined neural variability from another state-related perspective: the covariance of neural activity across brain areas is informative about behavioral states.[66,67]

Our approach identified novel latent states that are distinct from those associated with pupil diameter. Pupil diameter, arousal, and task performance have been reported to be positively correlated with each other[38,39,68–70] or sometimes exhibit an inverted-U-shaped correlation.[20,23,25,71,72] Pupil diameter fluctuations are sensitive to many factors: experimental setup (e.g., the light levels in the behavioral rig), behavioral task design (e.g., task difficulty), and physiological factors (e.g., heart rate, locomotion).[73,74] A possible explanation for why we did not observe a pupil-performance correlation is that our mice were placed on a platform instead of a running wheel, limiting the occurrence of high arousal states seen with locomotion.[20,25] Low arousal states may have likewise been limited by our approach: our inter-trial intervals were 1–2 s, instead of 3–7 s (typical of other studies). Longer inter-trial intervals allow pupil diameter to recover from reward-induced pupil dilations and for arousal to decrease before the next trial. By eliminating very high and very low arousal states, our approach may have unmasked movement signatures of novel cognitive states. Moreover, several recent studies have reported that spontaneous movement may be a better indicator of cortical excitation level than pupil size. Spontaneous facial movement and locomotion are coupled with high variance in cholinergic and noradrenergic signals in the dorsal cortex.[75,76] Together with our results, these findings raise the possibility that movement is informative about different axes of latent state.

More work is needed to understand the relationship between movements, internal states, and neural activity. First, one limitation of this study is that isolating sensory responses within the auditory cortex is difficult with our wide-field imaging preparation. Future work that examines state-dependent responses in the primary visual cortex during visual decisions will definitively determine whether disengagement alters sensory encoding. Second, single-cell recordings would address whether the change in the variability of neural activity is uniform across neurons and further dissect the underlying neuronal mechanisms more precisely. Third, the relatively low neural variance explained by task variables (Figure 3E, left) limits detecting task-variable-related neural differences across engagement states. A small difference might go undetected given the minimal variance captured by the model. Single-cell recording may also help to address this issue by identifying the single neurons with higher task-variable-related cvR^2 and testing whether this cvR^2 fluctuates with engagement. Fourth, our analyses focused on superficial structures accessible to wide-field imaging. Previous studies reported that deeper structures (e.g., basal forebrain, thalamus, locus coeruleus) are implicated in state-related regulation of neural activity through subcortical-cortical projections or neuromodulation.[71,77–81] An appealing idea is that the same modulatory inputs drive the movement pattern changes we observed. Future work on these structures could provide a mechanistic explanation for the observations we report here. Fifth, a manipulation experiment could establish a causal link between movements and internal state. Our approach cannot resolve whether TIM drives the animal to disengagement or instead reflects a disengaged state that was already initiated. However, this is a challenge given the difficulty of physically manipulating animals' movements and interpreting resulting changes in decision-making. As an alternative, targeted manipulation of state-modulating regions, such as the basal forebrain,[79,80] thalamus,[81,82] or locus coeruleus,[71,83] could be used to promote an engaged or disengaged state. We could then observe if TIM reflects such artificial state changes. Cell-type-specific measurements and manipulation are also needed and perhaps should extend to astrocytes as well as neurons.[34,71,84–87]

Previous observations demonstrated that the majority of neural activity is associated with movements but left unclear how such movements were related to cognition.[12,43,53] Our work uncovers that movements and cognition are intertwined: when animals become engaged in a task, movement patterns change, as does the trial-to-trial variability in neural activity. This points to a close link between movements and cognition and calls for new studies to understand how this link differs across species, areas, and behaviors.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by lead contact, Anne K. Churchland, Ph.D. (achurchland@mednet.ucla.edu).

Materials availability

This study did not generate new, unique reagents. Information about the transgenic mice, data collection equipment, and data analysis approaches is available from the lead contact upon request.

Data and code availability

The wide-field imaging data and corresponding behavioral videos used in this study are available at https://doi.org/10.25452/figshare.plus.21538458. A link to the data repository with a description of the behavioral and imaging data can be found at https://churchlandlab.dgsom.ucla.edu/code. All codes used to run the analyses and generate the figures in this manuscript are located at https://github.com/churchlandlab/TIM_state_manuscript/ (DOI: https://doi.org/10.5281/zenodo.15557838). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by U19NS123716, a Multi-university Research Grant from the Office of Naval Research (AW911NF-16-1-0368), a grant from the National Eye Institute (R01EY022979), a grant supporting the UCLA-Caltech Medical Scientist Training Program (1T32GM152342-01), and an NSF-NCS collaborative award (2219946). We thank Anup Khanal for assistance with animal handling and training. We also thank Michael Ryan, Felicia Davatolhagh, Lukas Oesch, Lillian Wilkins, and Letizia Ye for their helpful discussions and comments on the manuscript.

AUTHOR CONTRIBUTIONS

C.Y., M.D.M., and A.K.C. designed the study. C.Y., M.D.M., and G.R.-B. performed the analyses. C.Y., M.D.M., and A.K.C. wrote the manuscript with input from all authors. A.K. helped to pre-process behavioral data. The wide-field imaging and corresponding behavioral data (task 1) were collected by X.R. S., S.G., and S.M. at Cold Spring Harbor Laboratory. This dataset was first

published in 2023.[26] The pupil and behavioral data of task 2 were collected by J.C. The behavioral data of task 3 were collected by C.Y.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Animal information
  - General surgical procedures
- METHOD DETAILS
  - Auditory task for the mice used in widefield imaging (Task 1)
  - Visual task for the mice used in pupil-engagement analysis (Task 2)
  - Rat behavioral training (Task 3)
  - DeepLabCut tracking and motion energy measurement
  - Widefield imaging
  - Preprocessing of neural data
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - GLM-HMM model selection and state inference
  - Psychometric function fitting
  - Analysis of single trial latents
  - Linear encoding model
  - Task-Independent Movement (TIM) calculation
  - Pupil diameter measurement

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2025.06.001.

Received: August 16, 2024
Revised: February 4, 2025
Accepted: June 2, 2025
Published: July 1, 2025

REFERENCES

1. Najafi, F., and Churchland, A.K. (2018). Perceptual Decision-Making: A Field in the Midst of a Transformation. Neuron *100*, 453–462. https://doi.org/10.1016/j.neuron.2018.10.017.
2. Okazawa, G., and Kiani, R. (2023). Neural Mechanisms That Make Perceptual Decisions Flexible. Annu. Rev. Physiol. *85*, 191–215. https://doi.org/10.1146/annurev-physiol-031722-024731.
3. Runyan, C.A., Piasini, E., Panzeri, S., and Harvey, C.D. (2017). Distinct timescales of population coding across cortex. Nature *548*, 92–96. https://doi.org/10.1038/nature23020.
4. Wang, X.-J. (2012). Neural dynamics and circuit mechanisms of decision-making. Curr. Opin. Neurobiol. *22*, 1039–1046. https://doi.org/10.1016/j.conb.2012.08.006.
5. Deneve, S., Latham, P.E., and Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. Nat. Neurosci. *4*, 826–831. https://doi.org/10.1038/90541.
6. Shadlen, M.N., and Newsome, W.T. (1994). Noise, neural codes and cortical organization. Curr. Opin. Neurobiol. *4*, 569–579. https://doi.org/10.1016/0959-4388(94)90059-0.
7. Shadlen, M.N., and Newsome, W.T. (1998). The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. J. Neurosci. *18*, 3870–3896. https://doi.org/10.1523/JNEUROSCI.18-10-03870.1998.
8. Talluri, B.C., Kang, I., Lazere, A., Quinn, K.R., Kaliss, N., Yates, J.L., Butts, D.A., and Nienborg, H. (2023). Activity in primate visual cortex is minimally driven by spontaneous movements. Nat. Neurosci. *26*, 1953–1959. https://doi.org/10.1038/s41593-023-01459-5.
9. Yates, J.L., Coop, S.H., Sarch, G.H., Wu, R.-J., Butts, D.A., Rucci, M., and Mitchell, J.F. (2023). Detailed characterization of neural selectivity in free viewing primates. Nat. Commun. *14*, 3656. https://doi.org/10.1038/s41467-023-38564-9.
10. Raposo, D., Sheppard, J.P., Schrater, P.R., and Churchland, A.K. (2012). Multisensory Decision-Making in Rats and Humans. J. Neurosci. *32*, 3726–3735. https://doi.org/10.1523/JNEUROSCI.4998-11.2012.
11. Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. Science *340*, 95–98. https://doi.org/10.1126/science.1233912.
12. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. Nat. Neurosci. *22*, 1677–1686. https://doi.org/10.1038/s41593-019-0502-4.
13. Tremblay, S., Testard, C., DiTullio, R.W., Inchauspé, J., and Petrides, M. (2023). Neural cognitive signals during spontaneous movements in the macaque. Nat. Neurosci. *26*, 295–305. https://doi.org/10.1038/s41593-022-01220-4.
14. Miller, C.T., Gire, D., Hoke, K., Huk, A.C., Kelley, D., Leopold, D.A., Smear, M.C., Theunissen, F., Yartsev, M., and Niell, C.M. (2022). Natural behavior is the language of the brain. Curr. Biol. *32*, R482–R493. https://doi.org/10.1016/j.cub.2022.03.031.
15. Fu, O., Iwai, Y., Narukawa, M., Ishikawa, A.W., Ishii, K.K., Murata, K., Yoshimura, Y., Touhara, K., Misaka, T., Minokoshi, Y., et al. (2019). Hypothalamic neuronal circuits regulating hunger-induced taste modification. Nat. Commun. *10*, 4560. https://doi.org/10.1038/s41467-019-12478-x.
16. Melin, M.D., Khanal, A., Vasquez, M., Ryan, M.B., Churchland, A.K., and Couto, J. (2024). Large scale, simultaneous, chronic neural recordings from multiple brain areas. Preprint at bioRxiv. https://doi.org/10.1101/2023.12.22.572441.
17. Steinmetz, N.A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., et al. (2021). Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. Science *372*, eabf4588. https://doi.org/10.1126/science.abf4588.
18. Jun, J.J., Steinmetz, N.A., Siegle, J.H., Denman, D.J., Bauza, M., Barbarits, B., Lee, A.K., Anastassiou, C.A., Andrei, A., Aydın, Ç., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. Nature *551*, 232–236. https://doi.org/10.1038/nature24636.
19. Ashwood, Z.C., Roy, N.A., Stone, I.R., International Brain Laboratory, Urai, A.E., Churchland, A.K., Pouget, A., and Pillow, J.W. (2022). Mice alternate between discrete strategies during perceptual decision-making. Nat. Neurosci. *25*, 201–212. https://doi.org/10.1038/s41593-021-01007-z.
20. Hulsey, D., Zumwalt, K., Mazzucato, L., McCormick, D.A., and Jaramillo, S. (2024). Decision-making dynamics are predicted by arousal and unstructured movements. Cell Rep. *43*, 113709. https://doi.org/10.1016/j.celrep.2024.113709.
21. Bandi, A.C., and Runyan, C.A. (2024). Different state-dependence of population codes across cortex. Preprint at bioRxiv. https://doi.org/10.1101/2024.05.23.595581.
22. Bolkan, S.S., Stone, I.R., Pinto, L., Ashwood, Z.C., Iravedra Garcia, J.M., Herman, A.L., Singh, P., Bandi, A., Cox, J., Zimmerman, C.A., et al. (2022). Opponent control of behavior by dorsomedial striatal pathways depends on task demands and internal state. Nat. Neurosci. *25*, 345–357. https://doi.org/10.1038/s41593-022-01021-9.
23. McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., Tolias, A.S., Cardin, J.A., and McCormick, D.A. (2015). Waking State: Rapid Variations Modulate Neural and Behavioral Responses. Neuron *87*, 1143–1161. https://doi.org/10.1016/j.neuron.2015.09.012.

24. Vinck, M., Batista-Brito, R., Knoblich, U., and Cardin, J.A. (2015). Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. Neuron *86*, 740–754. https://doi.org/10.1016/j.neuron.2015.03.028.

25. McGinley, M.J., David, S.V., and McCormick, D.A. (2015). Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron *87*, 179–192. https://doi.org/10.1016/j.neuron.2015.05.038.

26. Musall, S., Sun, X.R., Mohan, H., An, X., Gluf, S., Li, S.-J., Drewes, R., Cravo, E., Lenzi, I., Yin, C., et al. (2023). Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. Nat. Neurosci. *26*, 495–505. https://doi.org/10.1038/s41593-022-01245-9.

27. Friedman, A., Hueske, E., Drammis, S.M., Toro Arana, S.E., Nelson, E.D., Carter, C.W., Delcasso, S., Rodriguez, R.X., Lutwak, H., DiMarco, K.S., et al. (2020). Striosomes Mediate Value-Based Learning Vulnerable in Age and a Huntington's Disease Model. Cell *183*, 918–934.e49. https://doi.org/10.1016/j.cell.2020.09.060.

28. Mazzucato, L., La Camera, G., and Fontanini, A. (2019). Expectation-induced modulation of metastable activity underlies faster coding of sensory stimuli. Nat. Neurosci. *22*, 787–796. https://doi.org/10.1038/s41593-019-0364-9.

29. Friedman, A., Slocum, A.F., Tyulmankov, D., Gibb, L.G., Altshuler, A., Ruangwises, S., Shi, Q., Toro Arana, S.E., Beck, D.W., Sholes, J.E.C., et al. (2016). Analysis of complex neural circuits with nonlinear multidimensional hidden state models. Proc. Natl. Acad. Sci. USA *113*, 6538–6543. https://doi.org/10.1073/pnas.1606280113.

30. Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., et al. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell *181*, 936–953.e20. https://doi.org/10.1016/j.cell.2020.04.007.

31. Marcos, E., Pani, P., Brunamonti, E., Deco, G., Ferraina, S., and Verschure, P. (2013). Neural Variability in Premotor Cortex Is Modulated by Trial History and Predicts Behavioral Performance. Neuron *78*, 249–255. https://doi.org/10.1016/j.neuron.2013.02.006.

32. Churchland, A.K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., and Shadlen, M.N. (2011). Variance as a Signature of Neural Computations during Decision Making. Neuron *69*, 818–831. https://doi.org/10.1016/j.neuron.2010.12.037.

33. Rabinowitz, N.C., Goris, R.L., Cohen, M., and Simoncelli, E.P. (2015). Attention stabilizes the shared gain of V4 populations. eLife *4*, e08998. https://doi.org/10.7554/eLife.08998.

34. Harris, K.D., and Shepherd, G.M.G. (2015). The neocortical circuit: themes and variations. Nat. Neurosci. *18*, 170–181. https://doi.org/10.1038/nn.3917.

35. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat. Neurosci. *21*, 1281–1289. https://doi.org/10.1038/s41593-018-0209-y.

36. Syeda, A., Zhong, L., Tung, R., Long, W., Pachitariu, M., and Stringer, C. (2024). Facemap: a framework for modeling neural activity based on orofacial tracking. Nat. Neurosci. *27*, 187–195. https://doi.org/10.1038/s41593-023-01490-6.

37. Harris, K.D. (2020). Nonsense correlations in neuroscience. Preprint at bioRxiv. https://doi.org/10.1101/2020.11.29.402719.

38. Van Kempen, J., Loughnane, G.M., Newman, D.P., Kelly, S.P., Thiele, A., O'Connell, R.G., and Bellgrove, M.A. (2019). Behavioural and neural signatures of perceptual decision-making are modulated by pupil-linked arousal. eLife *8*, e42541. https://doi.org/10.7554/eLife.42541.

39. Urai, A.E., Braun, A., and Donner, T.H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat. Commun. *8*, 14637. https://doi.org/10.1038/ncomms14637.

40. Joshi, S., Li, Y., Kalwani, R.M., and Gold, J.I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron *89*, 221–234. https://doi.org/10.1016/j.neuron.2015.11.028.

41. Aydın, Ç., Couto, J., Giugliano, M., Farrow, K., and Bonin, V. (2018). Locomotion modulates specific functional cell types in the mouse visual thalamus. Nat. Commun. *9*, 4882. https://doi.org/10.1038/s41467-018-06780-3.

42. Wang, Z.A., Chen, S., Liu, Y., Liu, D., Svoboda, K., Li, N., and Druckmann, S. (2023). Not everything, not everywhere, not all at once: a study of brain-wide encoding of movement. Preprint at bioRxiv. https://doi.org/10.1101/2023.06.08.544257.

43. Salkoff, D.B., Zagha, E., McCarthy, E., and McCormick, D.A. (2020). Movement and performance explain widespread cortical activity in a visual detection task. Cereb. Cortex *30*, 421–437. https://doi.org/10.1093/cercor/bhz206.

44. Gershman, S.J., Assad, J.A., Datta, S.R., Linderman, S.W., Sabatini, B.L., Uchida, N., and Wilbrecht, L. (2024). Explaining dopamine through prediction errors and beyond. Nat. Neurosci. *27*, 1645–1655. https://doi.org/10.1038/s41593-024-01705-4.

45. Markowitz, J.E., Gillis, W.F., Jay, M., Wood, J., Harris, R.W., Cieszkowski, R., Scott, R., Brann, D., Koveal, D., Kula, T., et al. (2023). Spontaneous behaviour is structured by reinforcement without explicit reward. Nature *614*, 108–117. https://doi.org/10.1038/s41586-022-05611-2.

46. Hughes, R.N., Bakhurin, K.I., Petter, E.A., Watson, G.D.R., Kim, N., Friedman, A.D., and Yin, H.H. (2020). Ventral Tegmental Dopamine Neurons Control the Impulse Vector during Motivated Behavior. Curr. Biol. *30*, 2681–2694.e5. https://doi.org/10.1016/j.cub.2020.05.003.

47. Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., Tank, D.W., et al. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature *570*, 509–513. https://doi.org/10.1038/s41586-019-1261-9.

48. Dhawale, A.K., Smith, M.A., and Ölveczky, B.P. (2017). The Role of Variability in Motor Learning. Annu. Rev. Neurosci. *40*, 479–498. https://doi.org/10.1146/annurev-neuro-072116-031548.

49. Wu, H.G., Miyamoto, Y.R., Gonzalez Castro, L.N.G., Ölveczky, B.P., and Smith, M.A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. Nat. Neurosci. *17*, 312–321. https://doi.org/10.1038/nn.3616.

50. Horrocks, E.A.B., Rodrigues, F.R., and Saleem, A.B. (2024). Flexible neural population dynamics govern the speed and stability of sensory encoding in mouse visual cortex. Nat. Commun. *15*, 6415. https://doi.org/10.1038/s41467-024-50563-y.

51. Christensen, A.J., and Pillow, J.W. (2022). Reduced neural activity but improved coding in rodent higher-order visual cortex during locomotion. Nat. Commun. *13*, 1676. https://doi.org/10.1038/s41467-022-29200-z.

52. Bennett, C., Arroyo, S., and Hestrin, S. (2013). Subthreshold Mechanisms Underlying State-Dependent Modulation of Visual Responses. Neuron *80*, 350–357. https://doi.org/10.1016/j.neuron.2013.08.007.

53. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brain-wide activity. Science *364*, 255. https://doi.org/10.1126/science.aav7893.

54. Hasnain, M.A., Birnbaum, J.E., Nunez, J.L.U., Hartman, E.K., Chandrasekaran, C., and Economo, M.N. (2024). Separating cognitive and motor processes in the behaving mouse. Preprint at bioRxiv. https://doi.org/10.1101/2023.08.23.554474.

55. Panzeri, S., Moroni, M., Safaai, H., and Harvey, C.D. (2022). The structures and functions of correlations in neural population codes. Nat. Rev. Neurosci. *23*, 551–567. https://doi.org/10.1038/s41583-022-00606-4.

56. Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. Nat. Neurosci. *14*, 811–819. https://doi.org/10.1038/nn.2842.

57. Cowley, B.R., Snyder, A.C., Acar, K., Williamson, R.C., Yu, B.M., and Smith, M.A. (2020). Slow Drift of Neural Activity as a Signature of

Impulsivity in Macaque Visual and Prefrontal Cortex. Neuron *108*, 551–567.e8. https://doi.org/10.1016/j.neuron.2020.07.021.

58. Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M. R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B. B., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat. Neurosci. *13*, 369–378. https://doi.org/10.1038/nn.2501.
59. Cohen, M.R., and Maunsell, J.H.R. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. *12*, 1594–1600. https://doi.org/10.1038/nn.2439.
60. Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2007). Differential Attention-Dependent Response Modulation across Cell Classes in Macaque Visual Area V4. Neuron *55*, 131–141. https://doi.org/10.1016/j.neuron.2007.06.018.
61. Ni, A.M., Ruff, D.A., Alberts, J.J., Symmonds, J., and Cohen, M.R. (2018). Learning and attention reveal a general relationship between population activity and behavior. Science *359*, 463–465. https://doi.org/10.1126/science.aao0284.
62. Liska, J.P., Rowley, D.P., Nguyen, T.T.K., Muthmann, J.-O., Butts, D.A., Yates, J.L., and Huk, A.C. (2022). Running modulates primate and rodent visual cortex via common mechanism but quantitatively distinct implementation. eLife *12*, RP87736. https://doi.org/10.1101/2022.06.13.495712.
63. Kang, I., Talluri, B.C., Yates, J.L., Niell, C.M., and Nienborg, H. (2025). Is the impact of spontaneous movements on early visual cortex species specific? Trends Neurosci. *48*, 7–21. https://doi.org/10.1016/j.tins.2024.11.006.
64. Haimerl, C., Ruff, D.A., Cohen, M.R., Savin, C., and Simoncelli, E.P. (2023). Targeted V1 comodulation supports task-adaptive sensory decisions. Nat. Commun. *14*, 7879. https://doi.org/10.1038/s41467-023-43432-7.
65. Huang, C., Ruff, D.A., Pyle, R., Rosenbaum, R., Cohen, M.R., and Doiron, B. (2019). Circuit Models of Low-Dimensional Shared Variability in Cortical Networks. Neuron *101*, 337–348.e4. https://doi.org/10.1016/j.neuron.2018.11.034.
66. Benisty, H., Barson, D., Moberly, A.H., Lohani, S., Tang, L., Coifman, R.R., Crair, M.C., Mishne, G., Cardin, J.A., and Higley, M.J. (2024). Rapid fluctuations in functional connectivity of cortical networks encode spontaneous behavior. Nat. Neurosci. *27*, 148–158. https://doi.org/10.1038/s41593-023-01498-y.
67. Shahsavarani, S., Thibodeaux, D.N., Xu, W., Kim, S.H., Lodgher, F., Nwokeabia, C., Cambareri, M., Yagielski, A.J., Zhao, H.T., Handwerker, D.A., et al. (2023). Cortex-wide neural dynamics predict behavioral states and provide a neural basis for resting-state dynamic functional connectivity. Cell Rep. *42*, 112527. https://doi.org/10.1016/j.celrep.2023.112527.
68. Russell, L.E., Fişek, M., Yang, Z., Tan, L.P., Packer, A.M., Dalgleish, H.W. P., Chettih, S.N., Harvey, C.D., and Häusser, M. (2024). The influence of cortical activity on perception depends on behavioral state and sensory context. Nat. Commun. *15*, 2456. https://doi.org/10.1038/s41467-024-46484-5.
69. Nassar, M.R., Rumsey, K.M., Wilson, R.C., Parikh, K., Heasly, B., and Gold, J.I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. Nat. Neurosci. *15*, 1040–1046. https://doi.org/10.1038/nn.3130.
70. Ebitz, R.B., and Platt, M.L. (2015). Neuronal Activity in Primate Dorsal Anterior Cingulate Cortex Signals Task Conflict and Predicts Adjustments in Pupil-Linked Arousal. Neuron *85*, 628–640. https://doi.org/10.1016/j.neuron.2014.12.053.
71. Aston-Jones, G., and Cohen, J.D. (2005). AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance. Annu. Rev. Neurosci. *28*, 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709.
72. Schriver, B.J., Bagdasarov, S., and Wang, Q. (2018). Pupil-linked arousal modulates behavior in rats performing a whisker deflection direction discrimination task. J. Neurophysiol. *120*, 1655–1670. https://doi.org/10.1152/jn.00290.2018.
73. Burlingham, C.S., Mirbagheri, S., and Heeger, D.J. (2022). A unified model of the task-evoked pupil response. Sci. Adv. *8*, eabi9979. https://doi.org/10.1126/sciadv.abi9979.
74. Wang, C.-A., Baird, T., Huang, J., Coutinho, J.D., Brien, D.C., and Munoz, D.P. (2018). Arousal Effects on Pupil Size, Heart Rate, and Skin Conductance in an Emotional Face Task. Front. Neurol. *9*, 1029. https://doi.org/10.3389/fneur.2018.01029.
75. Lohani, S., Moberly, A.H., Benisty, H., Landa, B., Jing, M., Li, Y., Higley, M. J., and Cardin, J.A. (2022). Spatiotemporally heterogeneous coordination of cholinergic and neocortical activity. Nat. Neurosci. *25*, 1706–1713. https://doi.org/10.1038/s41593-022-01202-6.
76. Collins, L., Francis, J., Emanuel, B., and McCormick, D.A. (2023). Cholinergic and noradrenergic axonal activity contains a behavioral-state signal that is coordinated across the dorsal cortex. eLife *12*, e81826. https://doi.org/10.7554/eLife.81826.
77. Roth, M.M., Dahmen, J.C., Muir, D.R., Imhof, F., Martini, F.J., and Hofer, S. B. (2016). Thalamic nuclei convey diverse contextual information to layer 1 of visual cortex. Nat. Neurosci. *19*, 299–307. https://doi.org/10.1038/nn.4197.
78. Sara, S.J., and Bouret, S. (2012). Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. Neuron *76*, 130–141. https://doi.org/10.1016/j.neuron.2012.09.011.
79. Pinto, L., Goard, M.J., Estandian, D., Xu, M., Kwan, A.C., Lee, S.-H., Harrison, T.C., Feng, G., and Dan, Y. (2013). Fast modulation of visual perception by basal forebrain cholinergic neurons. Nat. Neurosci. *16*, 1857–1863. https://doi.org/10.1038/nn.3552.
80. Zhu, F., Elnozahy, S., Lawlor, J., and Kuchibhotla, K.V. (2023). The cholinergic basal forebrain provides a parallel channel for state-dependent sensory signaling to auditory cortex. Nat. Neurosci. *26*, 810–819. https://doi.org/10.1038/s41593-023-01289-5.
81. Neske, G.T., and Cardin, J.A. (2023). Transthalamic input to higher-order cortex selectively conveys state information. Preprint at bioRxiv. https://doi.org/10.1101/2023.10.08.561424.
82. Yang, M., Keller, D., Dobolyi, A., and Valtcheva, S. (2025). The lateral thalamus: a bridge between multisensory processing and naturalistic behaviors. Trends Neurosci. *48*, 33–46. https://doi.org/10.1016/j.tins.2024.11.005.
83. Jordan, R. (2024). The locus coeruleus as a global model failure system. Trends Neurosci. *47*, 92–105. https://doi.org/10.1016/j.tins.2023.11.006.
84. McCormick, D.A., Nestvogel, D.B., and He, B.J. (2020). Neuromodulation of Brain State and Behavior. Annu. Rev. Neurosci. *43*, 391–415. https://doi.org/10.1146/annurev-neuro-100219-105424.
85. Reimer, J., McGinley, M.J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D.A., and Tolias, A.S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat. Commun. *7*, 13289. https://doi.org/10.1038/ncomms13289.
86. Fu, Y., Tucciarone, J.M., Espinosa, J.S., Sheng, N., Darcy, D.P., Nicoll, R. A., Huang, Z.J., and Stryker, M.P. (2014). A Cortical Circuit for Gain Control by Behavioral State. Cell *156*, 1139–1152. https://doi.org/10.1016/j.cell.2014.01.050.
87. Vaidyanathan, T.V., Collard, M., Yokoyama, S., Reitman, M.E., and Poskanzer, K.E. (2021). Cortical astrocytes independently regulate sleep depth and duration via separate GPCR pathways. eLife *10*, e63329. https://doi.org/10.7554/eLife.63329.
88. Couto, J., Musall, S., Sun, X.R., Khanal, A., Gluf, S., Saxena, S., Kinsella, I., Abe, T., Cunningham, J.P., Paninski, L., et al. (2021). Chronic, cortex-wide imaging of specific cell populations during behavior. Nat. Protoc. *16*, 3241–3263. https://doi.org/10.1038/s41596-021-00527-z.
89. Linderman, S. (2023). GitHub repository. https://github.com/lindermanlab/ssm.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Mouse: Emx1-IRES-Cre: Emx1$^{tm1(cre)Krj}$ (Task 1, Figures 1, 2, 3, 4, 5, 6, and 7)	The Jackson Laboratory	JAX#005628
Mouse: Fezf2-2A-CreER: Fezf2$^{tm1.1(cre/ERT2)Zjh}$ (Task 1, Figures S7 and S10)	The Jackson Laboratory	JAX#036296
Mouse: Ai162(TIT2L-GC6s-ICL-tTA2)-D (Ai162D): Igs7$^{tm162.1(tetO-GCaMP6s,CAG-tTA2)Hze}$ (Task 1, Figure S10)	H. Zeng, Allen Institute for Brain Science	JAX#031562
Mouse: PlexinD1-2A-CreER: Plxnd1$^{tm2.1(flpo)Zjh}$ (Task 1, Figure S10)	The Jackson Laboratory	JAX#036295
Mouse: B6.Cg-Tg(Camk2a-tTA)1Mmay/DboJ (Task 2, Figures 6 and 7)	The Jackson Laboratory	JAX#007004
Mouse: C57BL/6J (Task 2, Figures 6 and 7)	The Jackson Laboratory	JAX#000664
Mouse: B6;DBA-Tg(tetO-GCaMP6s)2Niell/J (Task 2, Figures 6 and 7)	The Jackson Laboratory	JAX#024742
Long-Evans rat (Task 3, Figures 7 and S13)	Charles River Laboratories	006
Software and algorithms		
MATLAB R2018a/R2021b	MathWorks	SCR_001622
Python 3.9.13	Python Software Foundation	SCR_008394
DeepLabCut 2.1.5.2/2.3.4	Mathis Group & Mathis Lab at EPFL	SCR_021391
SSM 0.0.1	Scott Linderman Lab, Stanford University	https://github.com/lindermanlab/ssm
Customized MATLAB & Python code	Authors of this paper	https://doi.org/10.5281/zenodo.15557838

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animal information

All surgical and behavioral procedures adhered to the guidelines established by the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee of Cold Spring Harbor Laboratory and the University of California, Los Angeles David Geffen School of Medicine. We used 8-25 week old male and female mice and 23-62 week old male Long-Evans rats. Detailed information on the animal source is provided in the key resources table. No statistical methods were used to pre-determine sample sizes. Sample sizes are similar to previous publications. Mouse strains were acquired from the Jackson Laboratory, Allen Brain Institute, or generated at Cold Spring Harbor Laboratory. The mouse room had a relative humidity of 30-70%, and room temperature ranging from 69-78°F. To avoid potential aberrant cortical activity patterns, EMX mice were on a doxycycline (DOX)-containing diet, preventing GCaMP6s expression until they were 6 weeks or older.

General surgical procedures

Surgeries were performed under 1-2% isoflurane in oxygen anesthesia. After induction of anesthesia, 1.2 mg/kg meloxicam was injected subcutaneously and sterile lidocaine ointment was applied topically to the skin incision site. After making a midline cranial incision, the skin was retracted laterally and fixed in position with tissue adhesive (Vetbond, 3M). We then built an outer wall using dental cement (C&B Metabond, Parkell; Ortho-Jet, Lang Dental) along the lateral edge of the dorsal cranium (frontal and parietal bones). A custom titanium skull post was then attached to the dental cement. For skull clearing, the skull was thoroughly cleaned followed by the application of a thin layer of cyanoacrylate (Zap-A-Gap CA+, Pacer technology). Mice were allowed to recover for 5-7 days after the skull clearing before starting the data collection.

METHOD DETAILS

Auditory task for the mice used in widefield imaging (Task 1)

Data used here are from a previously published study.[26] The behavioral setup was controlled with a microcontroller-based (Arduino Due) finite state machine (Bpod r0.5, Sanworks) using custom Matlab code (2015b, Mathworks) running on a Linux PC. Servo motors (Turnigy TGY-306G-HV) and touch sensors were controlled by microcontrollers (Teensy 3.2, PJRC) running custom code. Mice were trained on a delayed, spatial discrimination task. Mice initiated trials by placing their hands on at least one of the two handles, which were mounted on servo motors that rotated out of reach during the inter-trial period. To initiate trials, animals placed their hands on the handles and, after a variable duration of 0.25-0.75s of continuous contact, the auditory stimulus was presented. Auditory stimuli consisted of a sequence of Poisson-distributed, 3-ms long auditory click sounds, presented from either a left and/or right speaker for a variable duration between 1 and 2 s. The stimulus epoch was followed by a variable delay (0 to 0.5 s), after which servo motors moved two lick spouts close to the animal's mouth. If the animal licked twice on the side where more sensory events were presented, a drop of water reward was dispensed. After a spout was licked twice, the contralateral spout moved out of reach to force the animal to commit to its decision. The water volume rewarded per trial (typically 1.5 to 3 μL) was constant within a single session but was sometimes adjusted daily based on the animal's body weight. 5 mice (out of 9) in Figure S10 were used in an optogenetic inactivation experiment based on the same task.[26] To ensure the inactivation always covered the same time window, the durations of stimulus and delay epochs for these mice were fixed to 1.0s and 0.5s respectively. All optogenetic inactivation trials were removed before calculating TIM in case optogenetic manipulation changes the movement pattern.

Data was collected from multiple sensors in the behavioral setup. Touch sensors using a grounding circuit on handles and lick spouts detected contact with the animal's hands and tongue, respectively. A piezo sensor (1740, Adafruit LLC) below the animal's trunk was used for monitoring body and hindlimb movements. Two webcams (C920 and B920, Logitech) were positioned to capture the animal's face/upper body (lateral view) and the ventral surface of the body (bottom view). The frame rate of both cameras was 30 Hz. The pupil diameter data in Figures 6A and 6C was extracted from the lateral videos with customized MATLAB code.

Trained mice were housed in groups of two or more under reverse light cycle (12-hour dark and 12-hour light) and trained during their active dark cycle. Animals were trained over the course of approximately 30-60 days.

Visual task for the mice used in pupil-engagement analysis (Task 2)

Mice were trained in a visual spatial-temporal discrimination decision making task to monitor pupil dynamics during task engagement. Visual stimuli timing was equivalent to Task 1, described above.[26] Stimuli were presented in a screen (LG LP097QX1, Adafruit) placed 10cm in front of the mouse covering 55 to -55 degrees, calibrated to approximately 50 lux. The eye was recorded through a camera placed at 70 degrees, in the mouse visual field, and 14cm away. The pupil was illuminated with an 850nm LED (IR30, CMVision) and recorded at 30-60Hz using a monochrome camera (CM3-U3-13Y3M-CS, FLIR) equipped with a 12mm lens (NVM-12M23, Navitar) and spacer (ACC-01-5004, FLIR). Custom hardware and software controlled the behavioral task and synchronization with the camera. A micro-controller (Teensy 4.0, PJRC) recorded task events and mouse licking, controlled task timing and emitted a TTL pulse on stimulus presentation that was connected to the general purpose pins of the camera. The status of GPIO was recorded for each frame to recover the precise timing of the stimulus. Data was compressed online using an Nvidia hardware video encoder (through FFMPEG) and stored to disk with custom software (https://github.com/jcouto/labcams).

Rat behavioral training (Task 3)

Freely moving rats were trained to do a 2-choice auditory/visual decision task based on a task developed previously.[10] Experiments took place in dark, sound-proof rigs (Industrial Acoustics, Bronx, NY). Trials with visual, auditory, and multisensory stimuli were interleaved. In the behavior setup (composed of one speaker, one LED board, and three ports), the rat initiated the trial by putting its nose in the central port. Next, the speaker and/or LED board started to play the stimulus sequence (sound clicks, white light flashes, or both). The number of stimulus events, presented over 1000 ms, ranged from 7 to 17. The decision boundary (12 click/flash stimuli) was abstract and learned with experience by the animal. When the number of stimulus events was above 12, right choices were rewarded; when the number was below 12, left choices were rewarded. The rewarded sides of the trials with exactly 12 stimuli were assigned randomly. The duration of the inter-event interval followed a Poisson distribution. The stimulus window was followed by a short delay period. The duration of the delay period was randomly assigned from an exponential distribution (mean = 0.12 s). After this delay period, an auditory go cue informed the animal that it could leave the central port and make the choice. After the rat poked its nose into the left or right port to report the choice, a reward (24 μL water) or a punishing sound (2.0s long sharp noise) was delivered based on whether the choice was correct. A 2.0 s long white noise sound was delivered if the rat left the central port too early (i.e., during the stimulus or delay period). A webcam (CM3-U3-13Y3M-CS 1/2" Chameleon®3 Monochrome Camera) was used to monitor animal movements. The camera was positioned above the animal and its frame rate was 80Hz.

DeepLabCut tracking and motion energy measurement

To quantify the movement of different body parts, we employed DeepLabCut (version 2.1.5.2 and 2.3.4), a neural-network-based motion tracking software, to track the movement of 27 body parts of the head-fixed mice performing the auditory decision-making task.[35] 14 body parts from the lateral videos and 13 parts from the bottom videos were labeled and tracked (Figure 4A). We trained

two DeepLabCut models (one for the lateral videos, and the other one for the bottom videos) for each of the mice. In each model, about 80-150 frames were extracted and labeled as the training input. The sample frames were selected from different sessions to make the tracking more robust across sessions. After training, we generated labeled videos to test the tracking quality, and kept retraining with new sample frames/refined labels until the tracking quality was visually acceptable (Videos S1 and S2). To further remove the outliers with poor tracking, we replaced all the values that are > 5 standard deviations away from their mean positions with NaNs.

We next measured the motion energy of DeepLabCut-tracked body parts in the stimulus and delay epochs. Motion energy was defined as the cumulative position change of each body part over frames. Because the lengths of stimulus and delay epochs are variable, the number of trials with the longest stimulus/delay epochs may be small. We omitted all frames that were longer than 70% of the longest stimulus/delay epoch to reduce the noise induced by low trial number. For trials that were shorter than this duration, timepoints after the trial ended were represented with NaNs. We standardized the motion energy of different body parts separately and then averaged all the frames and body parts in each trial together to get a single motion energy value for each trial.

To directly compare the motion energy of engaged and disengaged states, the trials with the highest 20% P(engaged) values of each animal were designated as ''engaged trials''. Similarly, the trials with the lowest 20% P(engaged) values were designated as ''disengaged trials''. To relate P(engaged) to other behavioral metrics, we first smoothed both P(engaged) with a 50-trial long Gaussian-weighted moving average filter. We then calculated Pearson's correlation coefficient between the smoothed P(engaged) and the other behavioral metric using the MATLAB function *corrcoef.m*.

Widefield imaging

Widefield imaging was done as reported previously.[12,26,88] In brief, we used an inverted tandem-lens macroscope and an sCMOS camera (Edge 5.5, PCO). Imaging resolution was 640 x 540 pixels after data were spatially binned at 4x. This resulted in a spatial resolution of about 20 micrometers per pixel. A 525 nm band-pass filter (#86-963, Edmund optics) was used to isolate fluorescence signal. Data were acquired at a framerate of 30 Hz, with alternating blue (470 nm, M470L3, Thorlabs) and violet light (405 nm, M405L3, Thorlabs) delivered along the same excitation path. This enabled isolation of calcium-dependent signals from intrinsic signals, such as hemodynamic responses. The 405 nm violet excitation captures non-calcium dependent GCaMP fluorescence, which can thus be subtracted from the signal obtained with blue (470 nm) excitation. All subsequent analyses were based on this subtracted signal. All frames were rigidly aligned to the Allen CCF (Common Coordinate Framework) using four anatomical landmarks on the surface of the skull.

Preprocessing of neural data

Neural data were preprocessed as described previously.[26] Briefly, rigid-body registration was used to register frames to the median of the first trial. We then used singular value decomposition (SVD) to compute the top 200 spatial and temporal components of the imaging data. These components comprised at least 95% of the total variance over each recording, and this was done to reduce computational requirements for subsequent analyses.

Imaging data was then aligned to five trial periods: pre-trial, handle grab (trial initiation), stimulus, delay, and outcome. Alignment was required because the duration of trial epochs was randomized to reduce temporal correlation between variables.

QUANTIFICATION AND STATISTICAL ANALYSIS

GLM-HMM model selection and state inference

To assess the correlation between TIM and states of engagement, we fit a 3-state hidden Markov model with Bernoulli generalized linear model observations (GLM-HMM).[19] The model is described by a KxK transition matrix (where K represents the number of states) and a set of weights for each state ($w_c^{(k)}$, where c represents the corresponding Bernouli GLM input parameter and k represents the state). We used two input parameters to the model, stimulus and bias, in order to predict choice behavior. The model was trained using expectation-maximization (EM). Model hyperparameters were selected with a grid search, where hyperparameter performance was assessed via 10-fold cross validation. The parameters included in this grid search were n_states = [1,2,3,4,5,6], alpha = [1,2], and sigma = [.25, 0.5, 0.75, 1], for a total of 48 parameter combinations. Because EM is not guaranteed to converge at a global maximum likelihood, we ran EM 10 times for each fold of cross-validation and picked the model with the highest cross-validated performance. The EM fitting was performed as previously described[19] using the SSM Python package.[89] We used maximum a posterior (MAP) estimation to estimate the HMM transition matrix and GLM weights. Models were fit to the combined data from all mice because some mice lacked enough data to train individual models. Posterior state probabilities were then inferred via the forward-backward algorithm, also implemented within the SSM package. Trials were assigned to disengaged or engaged states by applying a probability threshold of 0.8. Thus, if the probability of a trial being from a particular state was greater than 0.8, then that trial was assigned to the corresponding state (Figure 2A). Within sessions, we sampled equal numbers of trials from each state. Sessions with fewer than 25 trials in either state were excluded from analysis.

Psychometric function fitting

Psychometric functions were modeled with a four-parameter cumulative Gaussian:

$$\psi(x;\ \mu,\sigma,\gamma,\lambda)\ =\ \phi(x;\mu,\sigma)(1\ -\ \lambda\ -\ \gamma)+\gamma$$

Where x describes the stimulus evidence (negative indicates more leftward evidence, positive indicates more rightward evidence), σ describes the inverse slope, μ describes the bias, γ and λ define the low and high lapse rates respectively. ϕ is the cumulative normal function.

These parameters were estimated using Nelder-Mead Optimization. Psychometric curves were computed from the same trials that were used to train the GLM-HMM. When fitting psychometric curves for the GLM-HMM states, data from all mice were pooled.

Analysis of single trial latents

To measure variability in single-trial neural latents, we reduced the dimensionality of population activity using the singular value decomposition (SVD). We then computed the average variance across time in engaged or disengaged trials (Figures S3A and S3B), the variance across all engaged and disengaged timepoints (Figure S3C), and the variance across trials (Figure S3D) for the top 3 SVD dimensions.

Linear encoding model

The linear encoding model was implemented as previously described.[12] In brief, the linear encoding model included task- and movement-related variables. These variables were assembled into a design matrix containing analog and kernel-based regressors (Figure 3A; Table S1). After assembling the design matrix, the model was fit using ridge regression (with MLE-based estimation of the ridge parameters).

Encoding models were fit to data within single sessions. We fit the encoding model twice for every session: first for trials assigned by the GLM-HMM to the engaged state and then for trials assigned by the GLM-HMM to the disengaged state. Because there were typically more trials in the engaged state, these were randomly downsampled so that the trial numbers in each state for a session were always matched. Sessions with fewer than 50 engaged and 50 disengaged trials were discarded to eliminate models fit to insufficient data.

To assess the task-aligned and task-independent contributions of uninstructed movements (Figure 3G), the task-independent contribution was calculated as the increase in explained variance when movement variables were added to the task-variable-only model, representing the unique, task-independent influence of the uninstructed movement regressors. The task-aligned contribution was then determined by subtracting the task-independent component from the total explained variance of a model containing only uninstructed movements.

Task-Independent Movement (TIM) calculation

Our new metric, TIM, quantifies the magnitude of movements that are independent of task events. This process requires 2 steps. First, we took 7 task variables: current trial's stimulus strength ($v_{stim(n)}$), choice ($v_{choice(n)}$), outcome ($v_{reward(n)}$), an interaction term of choice and outcome ($v_{interaction(n)}$); the previous trial's choice ($v_{choice(n-1)}$), outcome ($v_{reward(n-1)}$), and an interaction term of choice and outcome ($v_{interaction(n-1)}$). These 7 variables were used as regressors in a linear model trained to predict the position of each labeled body part (Equation 1; Figure S9A). At each video frame (t), regressions were performed to predict the x and y values of all body parts ($\widehat{x}(t)$ or $\widehat{y}(t)$). Given the association between engagement and task performance, we balanced the number of correct and incorrect trials before performing the regression to avoid inducing regression quality bias to our analysis. Linear regression was next performed with customized MATLAB code. The R2 values of regressions were highly variable across mice, indicating that there are significant individual differences in the level of moving stereotypically (Figure S9B).

$$\widehat{x}(t)\ =\ \boldsymbol{\beta}_{1t}v_{stim(n)}+\boldsymbol{\beta}_{2t}v_{choice(n)}+\boldsymbol{\beta}_{3t}v_{reward(n)}+\boldsymbol{\beta}_{4t}v_{interaction(n)}+\boldsymbol{\beta}_{5t}v_{choice(n-1)}+\boldsymbol{\beta}_{6t}v_{reward(n-1)}$$

$$+\ \boldsymbol{\beta}_{7t}v_{interaction(n-1)}+\boldsymbol{\beta}_{0t} \qquad \text{(Equation 1)}$$

$\widehat{y}(t)$ was calculated in the same way. Next, we obtained TIM by calculating the Euclidean distance between model-predicted positions and actual positions (Equation 2).

$$TIM(t)\ =\ \sqrt{(x(t)\ -\ \widehat{x}(t))^2+(y(t)\ -\ \widehat{y}(t))^2} \qquad \text{(Equation 2)}$$

Similar to the motion energy calculation, we restricted calculations of TIM to the stimulus and delay epochs for each body part, and omitted all frames that were longer than 70% of the longest stimulus/delay epoch. The trials with shorter stimulus/delay epochs were filled with NaNs. TIM from each body part was then standardized (z-scored) to avoid the result being dominated by few body parts. We then averaged TIM across all body parts and frames to get a single value for each trial (Figure 5B).

Pupil diameter measurement

The pupil diameter of the mice from which we collected widefield imaging data (Figure 6A) was extracted from the lateral videos with customized MATLAB code.[12] We first binarized all the frames and selected the continuous areas bigger than 10 pixels. These areas are considered potential pupil areas. The whiskers sometimes cut the pupil area into multiple pieces in the binarized images. Disconnected areas were merged together to remove this effect. We next used MATLAB function *regionprops.m* to detect the centroids in the merged areas and computed their diameters. Missing values (e.g. due to poor image processing and failed centroid detection) were then extrapolated using forward and reverse autoregressive fits from the remaining samples. Individual pupil diameter traces were then smoothed with a 10-frame long averaging filter. Additionally, data outside six standard deviations were assigned zero weight in smoothing.

For data presented in Figure 6C, pupil diameter was extracted using customized software (https://bitbucket.org/jpcouto/mptracker) based on OpenCV.[41] In brief, a region of interest around the pupil was manually selected and used to convert pixels to mm, assuming an eye diameter of 6mm from which ~80% is visible. The contrast of individual frames was equalized using adaptive histogram equalization. Opening and/or closing morphological operations were done to minimize whisker artifacts and a threshold was applied to the image. Blobs resembling an ellipse with center in the eye were then extracted and the contours fit with an ellipse. The diameter is taken as the diameter of a circle with the same area as the ellipse.

For all the analyses presented in Figure 6, the pupil was smoothed with a gaussian filter (50 trials long).