bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Spatially distributed and regionally unbound cellular resolution brain-wide processing loops in mice

Michael Schartner[1], Ari Liu[1,2], International Brain Laboratory[1], Ila Fiete[1,2]

[1]IBL; [2]MIT

Abstract Until recently, it has been possible to examine activity in the brain globally through regional averaging or locally at cellular resolution. These studies characterized regions as functionally homogeneous entities (e.g. V1 for extracting low-level visual features) or single neurons in a region as heterogeneously tuned (e.g. mixed selectivity). Here, we leverage the unprecedentedly dense IBL electrophysiological recordings in mouse brains during a sensorimotor decision-making task and computationally combine these to generate a global event-aligned high temporal precision cellular-resolution functional map. We find that individual neurons specialize in simple functions, and ordering neurons by functional similarity reveals at cellular and millisecond resolution the temporal arc of computations unfurling over the phases of a trial and across the brain, providing a way to interpret neural variance beyond sensory and motor responses. Functionally similar cells are highly dispersed across regions, and each region contains cells of nearly all response types. Functional tuning is barely predictable by cytoarchitectural boundaries or spatial coordinates, or vice versa. The primary regional distinction is between cortical and subcortical processing, with subcortical neurons carrying most of the internal computations of expert animals, from integration to decision making and movement initiation. At most, the brain exhibits regional structure in a distributional sense: the flatness of an area's distribution of different functional cell types is lowest in cortical regions, but there is no clear correlation between distributional specialization and anatomical measures of cortical hierarchy. Finally, the high temporal resolution of electrophysiological recordings and the inclusion of subcortical areas reveal the existence of a novel set of robust collective and distributed processing networks in the brain, which rapidly emerge and dissolve across phases of a trial. These results point toward the idea of spatially wide-ranging dynamical processing – winding through most brain regions and involving a subset of neurons in each – to carry out basic functions, and contrast with the idea that brain areas are functionally specialized units.

Introduction

From phrenology to cytoarchitectural brain partitioning, the single area–single function hypothesis has been a durable concept. In fact, there are clear specializations in neural processing, from the sensory periphery (e.g. visual receptive fields [1], whisker mapping [2], auditory frequency mapping [3]) to high-level associative regions (face recognition areas [4, 5], groups of brain regions processing various movie aspects [6]), while sub-mm-resolution recordings show differential participation of brain areas in functionally distinct tasks [7, 8, 9, 10, 11, 12, 13, 14].

Regional averaging-based methods or sparse single-cell recording studies have characterized brain areas as having a relatively homogeneous function (e.g. V1 for extracting low-level visual features in cats [15]). Moreover, specializations exist on very large scales where circuits have distinct architectures, such as cortical (recurrent) versus cerebellar (largely feedforward) processing [16].

More modern recordings have shown, by contrast, that cells in high-level cortical areas can have heterogeneous tuning to multiple variables, the idea of multiplexing or mixed selectivity, which has been argued to be functionally useful for various reasons [17, 18]. These findings raise two critical and open questions: How widespread is cellular functional heterogeneity – does it extend across the brain or is it limited to higher brain areas that are hypothesized to synthesize information from all sensory streams? Does the idea of single area–single function hold when examined without regional averaging? And how does one reconcile regional function assignments with heterogeneous neural responses?

Regional specialization is consistent with a primarily feedforward perspective on brain function: information arrives at a region, is processed locally (possibly involving within-region recurrence), then is passed to the next region and so on. On the other hand, the brain possesses massive recurrence in the form of not only local interactions but also reentrant feedback

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

from multiple regions to determine ongoing activity. In this view, even basic functions may be distributed across multiple coordinating brain areas (fig. 1f). We previously lacked tools to discern, at cellular resolution with broad coverage, the prevalence of widespread recurrent and feedback processing.

To answer the questions above, we examine an unprecedentedly large (approximately 6×10^4 neuron) dataset of single-cell resolution Neuropixels recordings across the mouse brain as mice perform a simple visually-driven response task. In this task, internal expectations shape perception and guide the sequence of perception, evidence integration, decision-making, and action [19], engaging the whole brain [20]. We map how task demands, variables, and processes interact, and how their representations and dynamics unfold across the mouse brain. With appropriate neural ordering of this dense brainwide dataset, we can nearly see thoughts writ in neural activity and trace their progression over time and across the brain. Consistent with recent reports[20, 21], direct input-driven responses involve a small fraction of neurons, movement correlates with a substantially larger fraction, but the majority of cells encode correlated internal signals across the brain, such as the binary choice taken by the animal.

We find that most single neurons have specialized and non-mixed selectivity for specific functions, but cytoarchitecturally defined areas have highly diverse populations of functional neuron types that barely predict and are barely predicted by electrophysiological cell type, nor cytoarchitectural boundaries, nor 3D location in the brain. Conversely, neurons belonging to a functional class are homogeneously distributed across the brain, as found earlier for single regions [22]. The idea that certain brain regions mostly perform certain functions, holds at most for primary sensory areas. Correlating cellular responses at the region-wide level, we find that regions form transient processing networks, with dynamically shifting coalitions across different phases of the task. This dynamical reallocation of regional processing provides a distinct view than afforded by assigning a function to each area. Finally, we find that the most prominent regional distinction is between cortical versus subcortical groups, with subcortical areas performing most of the critical cognitive functions of the task. In sum, our results suggest that distributed brain-wide circuits implement even simple functions.

Results

On each trial of the IBL task, mice were presented a visual stimulus on the left or right of a screen and they reported their perceptual decision about the stimulus by moving a wheel clockwise or counterclockwise (thus steering the stimulus to the center of the screen), Fig. 1a. They received a water reward (time-out) if the response was correct (incorrect). Following a short and variable-duration inter-trial period (rest), animals had to be still (quiescence) for at least 200 ms before the next trial could begin. Trials were presented in blocks (of variable numbers of trials) in which the probability of left versus right stimulus appearance was 0.8:0.2, alternating with blocks where this probability was flipped (0.2:0.8). Thus, animals could in principle learn this block structure, estimate the block side, and leverage this information to make more correct decisions. Well-trained animals tended to do so [19, 23].

Neural activity was recorded with up to two Neuropixel probes per well-trained mouse, over 139 animals, resulting in dense single-neuron resolution coverage of the left hemisphere. The resulting dataset comprises approximately 6×10^4 neurons with histologically reconstructed anatomical locations aligned to a standard brain Fig. 1a,b (for complete details on task, see [19] and on dataset, see [20]).

The time-course of summed brain-wide activity over the arc of perceptual decision making

Despite the variable nature and durations of the trials, behavior, and cells, and the acquisition of neurons from different animals, we could jointly leverage the internal temporal structure of each trial and define multiple distinct trial types to obtain standardized and aligned functional response vectors for each neuron, allowing us to pool and compare them as follows.

Each trial consisted of multiple "events": the times of stimulus, movement, and feedback onset, and the onset of the quiescence-dependent period. We windowed and aligned each neuron's activity in a trial to the trial events (rest period start, quiescent period start, stimulus onset, pre-movement onset, movement onset, and feedback onset), Fig. 1a. For each window, we defined a partition of trials into types. For instance, in the quiescence window, trials were partitioned into left and right block groups, to construct an average left block and average right block response per cell. We thus generated a peri event-time histogram (PETH) for left and right blocks for each cell. Similarly, in the stimulus window we partitioned trials into eight types

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(whether the stimulus, choice, and block sides were on the left or right), and constructed an average response of each cell in the stimulus window to each of the eight trial types.

We first consider how the full brain-wide neural population responds across conditions and trial events, by taking a simple grand average of all cellular responses, Fig. 1c. These curves of mean brain-wide responses in each trial window and across trial types reveal a rich first view of brain states and computations: brain-wide activity in the required quiescent period before stimuli can be presented is significantly suppressed relative to activity during simple rest. Further, activity during quiescence is strongly modified by internal expectation of a broader task context: each session begins with 90 trials that have no block side bias (50_b), and the quiescent interval activity during these trials is much lower than during trials in which there is a left or right bias. Next, there are at least two distinct activity phases during the stimulus onset phase, a rapid response peak that decays and a slower ramping phase. The response ramps divide into two groups: On top are the 4 PETH types with correct response (i.e. congruent stimulus side and choice for each block type) and below the 4 mistake conditions (2 incorrect combinations, one per block). Movement onset is preceded by a ramp to a threshold, and the largest mean response is evoked immediately following positive and negative feedback. Despite the very different nature of the positive and negative feedback (water versus time-out), and the greater amplitude of response to positive feedback, the neural response decays to both types of feedback are, surprisingly, nearly identical, although positive feedback activity displays an overlaid water licking-related oscillation (see Extended Data Fig.12 in [20] for examples). Thus, these trial- and cell- grand-average PETHs reveal brain-wide processing over the arc of preparation, stimulus processing, decision making, movement initiation, and feedback. They show the relatively small effect of stimulus perception (rapid sharp stimulus response) on brain activity, the gradual generation of strong internal activity via integration (ramping activity), decision thresholding (arrival of the distinct movement-aligned PETH responses to a similar value at movement time), huge effects of feedback processing, and modulation of inter-trial states based on context and internal expectation.

Single neuron responses during the task form interpretable functional clusters

Next, we turn to cellular-resolution responses. For each neuron we concatenated the above-defined PETHs and a few others (e.g. rest PETH, positive feedback PETH; list of all PETHs in table 1) into a long vector and z-scored the result to obtain that neuron's "feature vector". An annotated feature vector from an example cell is shown in Fig. 1d. Thus, we engineered a succinct, standardized, equal-length and equal-variance functional description for each cell, exploiting the IBL task structure.

To define functional response categories, we cluster the resulting cellular feature vector matrix ($N_c \times N_T$ dimensions; $N_c = 61473$ is the number of cells and $N_T = 2496$ ($t_{bin} = 125ms$) is the length of each feature vector) via k-means. Using 10-15 clusters generates readily interpretable and reproducibly defined categories, Fig. 1e (Fig.S6 for more cluster numbers).

Response components whose largest or narrowest peak is within a specific trial event are identified as associated to that event: for instance, the first response component (Fig. 1e, 1st line from top, dark red) exhibits significantly increased firing in both the stimulus and pre-movement periods, but the responses are more sharply peaked when aligned to movement onset, thus it is considered to be a choice- or motor-related response. Comparing curves with movement onset aligned peaks (1st, 5th, and 7th lines), two (1st, 5th) are accelerating toward the time of movement onset, indicating choice/movement initialization related activity, while the 7th exhibits non-accelerating linear ramps, indicative of an integrating mode. The stimulus responses (which putatively include responses to both the auditory tones and visual stimuli) are small compared to choice-, movement-, and feedback-related responses (cf. Fig. 1c), and thus do not show up prominently in the clustered responses but are present as narrow peaks aligned to stimulus onset (3rd line). Strikingly, four distinct mistake response components, separate for the different mistake contexts, are apparent (2nd, 9th, 11th, 12th lines; gray blocks in Fig. 1d designate mistake trials, in which stimulus and choice are on opposite sides): these components are highly activated in one of the four mistake trials. Their response profiles during the mistake trials are relatively flat and of similar amplitude when stimulus-aligned or movement-aligned.

All components but one (6th line) show suppressed responses in the 50_b interval, which is the inter-trial period of the block of 90 trials initiating each recording session in which the stimuli are presented with 50% probability on either side. Finally, one component (13th line) shows linear rate ramps during every trial when aligned to stimulus onset but not movement onset, and which does not appear to be contextually modulated by stimulus side, block side, or choice side. We might tentatively interpret

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

this component as a timer from stimulus onset. Our goal in forming these functional clusters is to investigate the anatomical distribution of functions across anatomically defined brain regions, and thus use our brain-wide recordings to address the question of whether basic functions are assigned to specific brain regions (Fig. 1f, left), or whether each basic function is carried out by loops and pathways that involve brain-wide coordination across most regions (Fig. 1f, right).

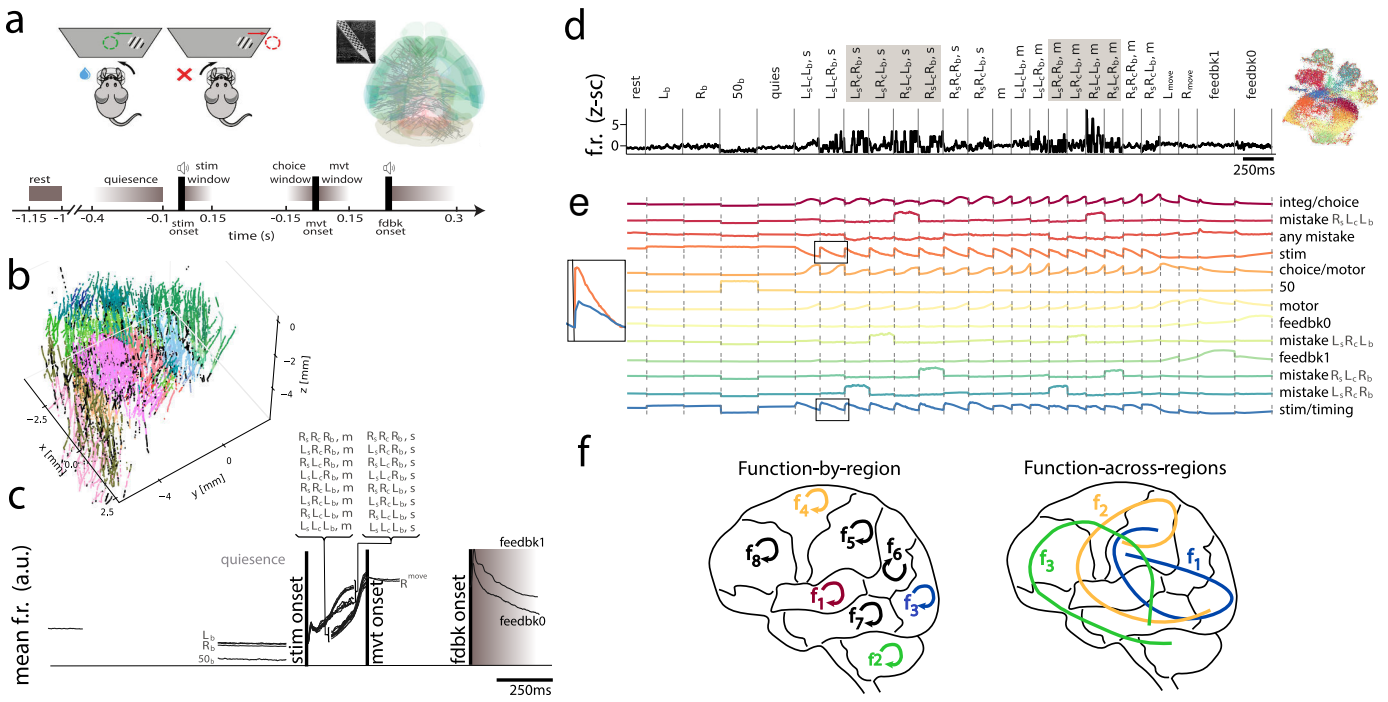


Figure 1. Task structure, functional clustering of brain-wide cellular responses, and hypothesis. a The IBL task: a mouse keeps still (quiescence period) to trigger a visual stimulus on the left or right (L_s or R_s). If it chooses and reports the correct stimulus side (L_c or R_c, respectively) by turning a wheel in the appropriate direction, it receives a reward (feedbk 1); otherwise it receives a timeout with white noise over the duration (feedbk 0). After a short rest period, the trial begins again. Stimulus and feedbk 1 onset are accompanied by an auditory cue. The probability the stimulus appears on the right side is elevated for a block of trials (R_b) of variable length; the probabilities then flip without an explicit cue so left stimuli become more probable (L_b). Neuropixel probe tracks across 139 mice that perform the task at or above expert criterion level densely tile the left hemisphere [20]. **b** The total yield for our analyses is 61473 neurons passing quality control across mice [20], rendered as dots colored by 260 anatomical "Beryl" regions. **c** Grand-average brain-wide neural responses aligned to key trial events, separated by trial type (8 types: left, right stim, choice, and block sides) within each event window: within a window, we average the responses of all cells across all trials of the same type (e.g. all feedbk 0 trials within the feedback window, or all trials with left stimulus, left choice, during a right block ($L_sL_cR_b$) in the stimulus window). These curves define peri-event time histograms (PETHs) averaged across all cells. See Tab.1 for details. PETHs between the stimulus onset and movement onset periods are denoted ($X_sX_cX_b, s$) for the stimulus window (stimulus aligned) and ($X_sX_cX_b, m$) for the choice window (movement aligned), where X designates L or R. The stimulus-aligned curves split towards their end in two groups, with the top 4 lines being stimulus side and choice congruent curves while the bottom ones are the 4 incongruent ones, i.e. "mistake" conditions. **d-e** Single-cell analyses. **d** Left: The *functional feature vector* of one neuron, consisting of 26 PETHs (all windows and all trial types in each window) concatenated then z-scored. Right: Each neuron's functional feature vector defines a point in a high-dimensional space, which we cluster using k-means in the high-dimensional space. Shown is a 2-dimensional umap projection of the feature vectors (each dot is one neuron, colored by its cluster assignment). **e** The average functional feature vector of all neurons in the 13 k-means clusters. These are interpretable (see horizontal labels; e.g. the "mistake $R_sL_cL_b$" cluster (red) consists of neurons that fire most during the stimulus window when the upcoming left choice is opposite the right stimulus, during the left block. Inset shows zoom of rapid stimulus responding clusters. **f** Schematic of two hypotheses for brain function: Brain computations are strongly regionalized (spatially modular); or functionally modular computations are performed in distributed multi-regional loops.

Low co-correspondence of function and anatomy

When we visualize the three-dimensional anatomical position of each recorded cell, colored by its functional k-means cluster, the result is a spatially homogeneous mix (Fig. 2a, left). By contrast, if we embed each cell into a two-dimensional space through UMAP on its feature vector, coloring by functional k-means clusters results in well-segregated color boundaries (Fig. 2a, right). Four "islands" in the functional UMAP embedding correspond to the four mistake conditions (cf. Fig. 1e). For easier visualization of functional-spatial correlations, we separately plot the spatial distributions of all neurons belonging to a given functional response (k-means) cluster, Fig. 2b. Here we find no discernible spatial localization of any functional cluster: each functional category is broadly represented across the brain.

Next, though each function might not be localized in terms of three-dimensional euclidean neighborhoods in the brain,

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

there might still be locality of function in terms of brain parcellations, for instance based on cytoarchitecture, which are spatially non-convex and extended regions. To investigate this possibility, we colored the UMAP embedding of feature vectors by different common parcellations of the brain (10 "Cosmos" regions, 260 "Beryl" regions, or 6 "layers", with the latter referring to the cortical layers with subcortical structures as another "layer"; see further methods). We find little to no order in the assignment of regions to functions, Fig. 2c, d, e. The primary discernible function-region correlation emerges from the Cosmos and Beryl colorings, which reveal a slight pink (hindbrain, midbrain)—green (isocortex, hippocampus) separation, and from the layer coloring, which reveal a matching and slight black non-black separation. We quantify this by smooth density plots of the umap space, separate for each Cosmos brain region, showing that regions with green cell colors - isocortex, hippocampal formation (HPF) - are distributed fairly homogeneously across the embedding space while cells from midbrain (MB) and hindbrain (HB), which have a pink color, are concentrated visibly on one side of the space, Fig. S5b.

We next determine whether the functional groups exhibit some spatial modularity in the brain, even though those spatial distributions might not be localized in three-dimensional space or by known brain parcellations. To do so, we use the per-function cellular spatial distribution data from Fig. 2b, spatially filter these to obtain kernel density estimates of the spatial distribution of each function, and compute the pairwise overlaps in these spatial distributions using cosine similarity (Methods). The result is a nearly uniform similarity matrix, Fig. 2f, suggesting that the distinct functions are not spatially modularized in the brain. As a control, the same analysis applied to the spatial distribution of neurons in Cosmos regions, from Fig. 2d (left), yields low similarity scores and thus clear modularity of the regions, Fig. 2g. We further controlled for coverage inhomogeneities, i.e. per-voxel divided each 3D density image by the density image of all cells prior to computing similarity scores for region pairs, finding no qualitative differences in the results, Fig. S16.

This is additionally supported when comparing distributions of centroid distances of each class. I.e. the Euclidean distance between the centroid of all cells and that of, say, a kmeans cluster is one point in the distribution for kmeans. We find 13 kmeans classes have mean 0.32 mm, std 0.25 mm; 204 Beryl classes: mean 2.97 mm, std 1.31; 12 Cosmos classes: mean 2.40 mm, std 1.14 mm and 8 layers classes: mean 1.71 mm, std 0.94 mm. I.e. the centroids of the kmeans classes are much closer to that of all cells, compared to the anatomical parcellations.

Finally, we augment these unsupervised approaches to relate function to anatomy via supervised decoding. We consider to what extent it is possible to classify a cell's anatomical region from its functional response vector. Anatomical decoding involved 204 Beryl regions (each containing at least 20 cells), 12 Cosmos regions (including undefined groups "root" and "void"), and canonical classes (7 classes of spatial brain regions historically defined as functionally distinct, Fig. 1a in [24]; listed in methods); "kmeans" (our 13 k-means clusters from earlier); and "layers" (neurons binned into cortical layers and thalamus) (see Methods for more details).

We split the cells repeatedly into train and test folds for classification of anatomical region from the neural response vectors, and show the test results for each fold, Fig. 2h. Across folds, decoding accuracy variance is small, as seen from the scatter in each point cloud. Test accuracies for all decoding tasks were significantly higher than chance level, determined by their respective shuffle controls (obtained by randomizing labels across cells), indicating that there is some non-zero information about anatomy in the functional responses. (The chance level varies across decoding tasks according to the number of classes, 6 layers < 7 canonical < 12 Cosmos < 13 k-means < 204 Beryl.) However, decoding accuracy is poor for all anatomical decoding tasks (the k-means categories are trivially decoded at accuracy 1).

For comparison, we consider how well another activity-related property of cells enables anatomical decoding: instead of the functional response vector for each neuron, we use 30 electrophysiological properties that are independent of the task, such as average firing rate, local field potential alpha power, and spike trough value (see Methods and [25] for details).

Overall, these "ephys features" yield slightly but significantly better anatomical decoding relative to controls than possible with the response vector (Fig. 2i); however, while significantly above chance, ephys features are visibly weaker predictive of the functional response class (Fig. 2i, red), suggesting that electrophysiological cell type is weakly correlated with functional response category.

In sum, in this unprecedented whole-brain dataset during a task with clear functional subtasks across phases of each trial, we find low to no regional localization of function in which regions are specialized to individual functions. We show that cells

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

190 performing all functions are homogeneously distributed across the brain, with at most a slight cortical-subcortical distinction.
191 This does not rule out the possibility that a given functional representation originates in an anatomically localized area, but it
192 indicates that even if so, that information is quickly distributed for use brainwide, so that the functional responses are nearly
193 homogeneously present across the brain.

194

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

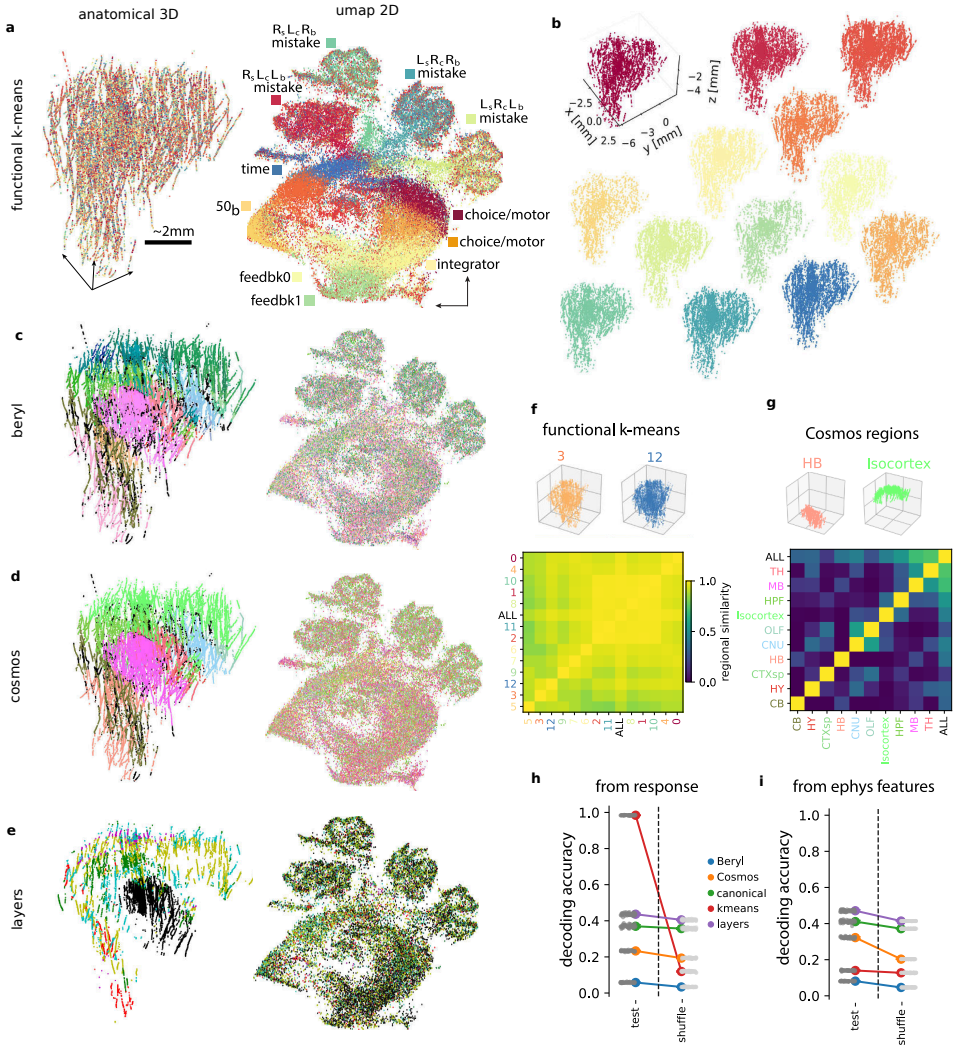


Figure 2. Non-correspondence between functional and anatomical clusters. a Anatomical 3D brain space reconstruction (1st column) and umap 2D functional embedding (2nd column) of all neurons. Coloring by 13 k-means clusters shows a homogeneous mixture in anatomical 3D but clear color boundaries in umap 2D, with functional descriptors (compare with Fig. 1e). **b** Disentangling the k-means clusters in 3D space shows their high visual similarity, each covering the whole left hemisphere homogeneously. **c-e** Conversely, the anatomical parcellations ("Beryl" **c**, "Cosmos" **d**, "layers" **e**) show the expected patterns in anatomical 3D clearly, while there are no clear color boundaries in umap 2D, apart from a weak pink (hindbrain, midbrain) green (isocortex, hippocampus) divide for "Beryl" and "Cosmos". **f** The high similarity of the 3D distribution of cells per k-means cluster is further quantified using the cosine similarity of normalised 3D density images, with two example clusters on top, dendrogram ordered. "ALL" indicates the group of all cells, having high similarity with all k-means clusters. **g** The same metric as in **f** applied to anatomical "Cosmos" brain regions describes them as highly non-similar in brain space distribution, in stark contrast to the k-means clusters. Compare with Fig. S16. **h** Decoding scores on cellular level of various target categories from concatenated PETHs ("from response"). Decoding accuracies are shown for multiple test folds (point clouds; STD smaller than mean icons) alongside scores with shuffled labels. All target categories can be decoded above chance from response (all solid lines indicate $p < 0.05$ of test/shuffle accuracies), although weakly, except k-means (red) with a near perfect decoding score. **i** Same analysis using different feature vectors. Instead of the concatenated PETHs, the input was 30 trial-structure-independent ephys features per cell, resulting in slightly higher decoding scores than from response for all targets but trivially k-means.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cell ordering by response similarity reveals the brainwide computational arc across time at cellular resolution

Our functional clustering procedure above assigned neurons to a function, however each neuron might be participating in multiple functions. Moreover, the analysis ignored detailed information on the temporal activity profiles of each neuron on every trial type. Thus, we sought a much finer-grained look at the brain-wide temporally-resolved activity evolution of all neurons over the course of all trial types.

To do so, we ordered all ~ 60k neurons according to the similarity in their response vectors via the Stringer rastermap[26]. The result is a striking view of mental computations writ in the activity states of individual neurons across trial phases, trial types, and time, Fig. 3a: There is clear block structure among neurons, seen as different temporally evolving patterns within trial types and across trial phases, and it is possible to see the temporal sweep of stimulus response (darkest portion of tornado-like pattern near cell count 10k), strong feedback-preferring neurons (largest dark region in lower right corner), error-encoding neurons (those firing only on trials where the choice is opposite the stimulus side, e.g. two block patterns at cell count 20k), and so on. To relate these rastermap blocks with our previously defined clusters (functional and anatomical), we overlay colors for each cell onto the rastermap, Fig. 3b-f. We find that previous functional clusters (Fig. 1e) form clear color blocks on the rastermap Fig. 3b, suggesting good consistency (as expected given that k-means is a component of the rastermap algorithm). The rastermap combined with the prior k-means clustering provides stronger interpretability of the clusters, which we will explore next in detail. Also consistent with our previous conclusions, there is a subcortical (hindbrain, midbrain)-cortical (isocortex, hippocampus) divide in neural response patterns (Fig. 3c-d), as seen in the pink-green color blocks, but no clear distinctions based on cortical layer and thalamus (Fig. 3e), while partially overlapping with the average firing rate per cell (Fig. 3f, darker is lower firing rate).

Further, we find that most temporal response patterns present in the rastermap are also present when we restrict the rastermap to cells from single brain regions, whether the regions are in cerebellum, thalamus, isocortex, hippocampus, or midbrain, Fig. 3g. Thus, the rastermap confirms and demonstrates in detail that individual Beryl areas contain cells showing essentially the full diversity of temporal, trial-phase and trial-type specific response patterns.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

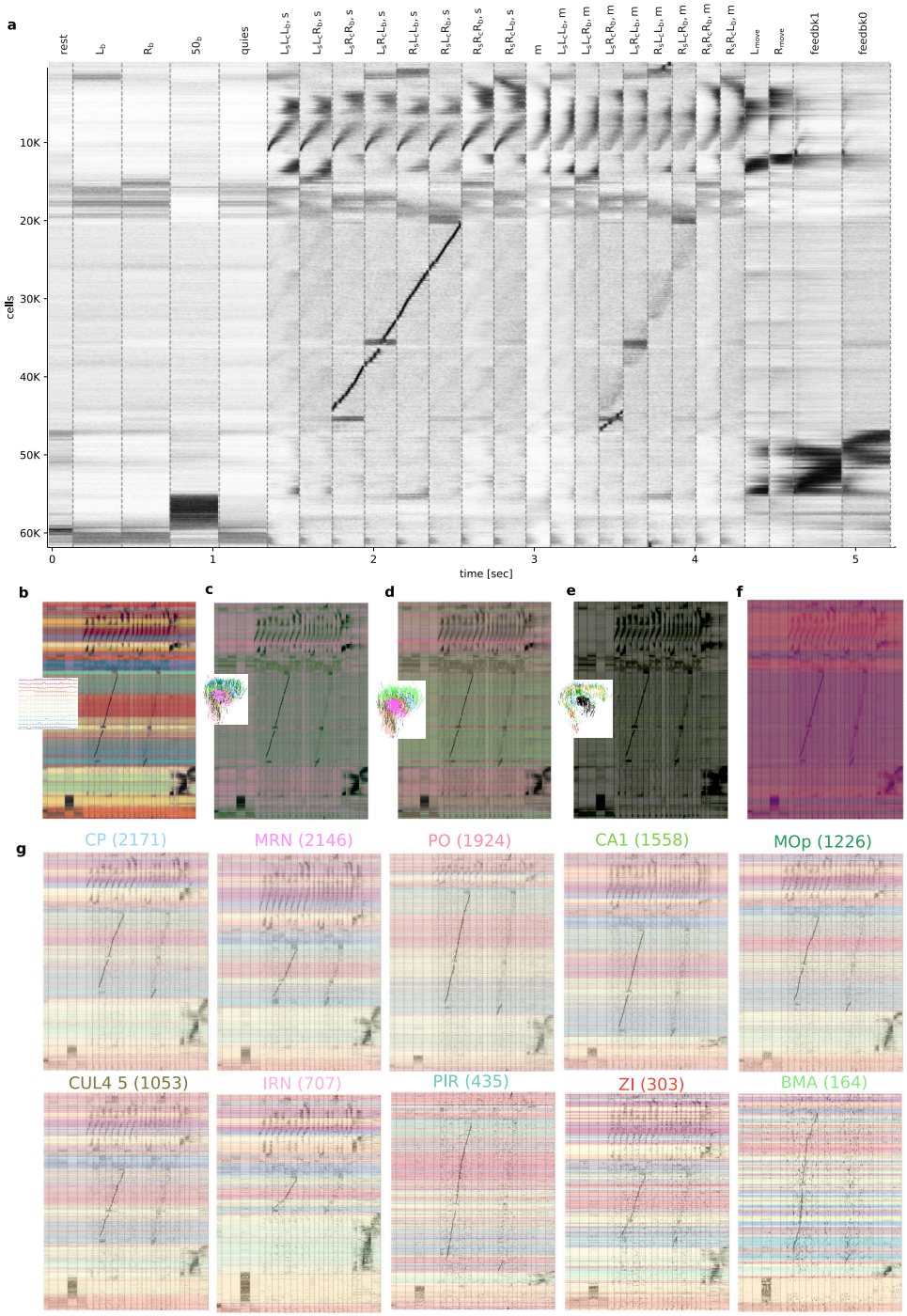


Figure 3. Ordering of neurons by similarity reveals functionally interpretable patterns. a Feature vectors of all ca. 60 k cells sorted by similarity via rastermap[26], revealing great diversity in activity and clear patterns. Black is high firing rate and white is low. Compare with example cell in Fig. 1d. **b** The same image colored by 4 different cell groupings: 13 k-means clusters with average feature vector of each cluster in inset, in line with rastermap. Note for example the cluster "mistake 1" (see also Fig. 1e) with two blocks of activity in cyan, top edge of inset. The coloring by the anatomical parcellations Beryl **c**, Cosmos **d** and layers **e** have a much less clear pattern and color matches. Beryl and Cosmos show a vague pink (hindbrain, midbrain) green (isocortex, hippocampus) color divide while layers seem homogeneously spread across, with thalamic (black) and neither layer nor thalamic cells (transparent) homogeneously painting the picture. **f** Firing rate of each cell (prior to z-scoring of feature vector), exponentiated by 0.1, darkest color is lowest. The majority of cells is sparsely firing. **g** Rastermap restricted to example regions, with neuron number in title next to acronym, k-means cluster as background colors. For each broad anatomical structure ("Cosmos") the region ("Beryl") with most neurons is shown. Most regions display most patterns seen for the full dataset in **a**, with smaller regions showing fewer pattern, however each contains cells of all k-means clusters. Also note that neuron number alone is not indicative of response diversity, e.g. region ZI having less neurons than PIR but larger overlap with the patterns seen for the full dataset.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

To advance the goal of understanding the timecourse of computations at cellular resolution and observing how various mental processes sweep through the brain over each trial type, we exploit the fine-grained temporal pattern interpretability afforded by the rastermap. We implemented a cross-validated procedure(S7), in which we sorted neurons using one set of randomly selected trials (50% of the total number of trials), then plotted their activities on the remaining (non-overlapping) set, Fig. 4a. This slightly restricted our dataset to cells with enough trials of each type (from $\sim$ 60K to $\sim$ 50K neurons). The cross-validation process is important because even random spiking across neurons can be ordered in a way that appears structured (see strong diagonal pattern in Fig. 3a, affecting mainly cells with low firing rate, which are more than half of all cells, Fig. 3f); the cross-validated rastermap only reveals temporal patterns that are reliably repeated across trials.

We can parse neurons into distinct functional groupings in the rastermap based on their temporally evolving activity patterns within each trial, their differential responses across trial phases and trial types, and differences in firing rates. After doing so, we identify a group of neurons as stimulus responding cells (black box, top Fig. 4a). Their activity is characterized by a narrow time-locked response at stimulus onset, regardless of trial type, and also at the start of the feedback interval where sensory stimuli (such as an auditory click) accompany feedback onset. Consistent with this characterization, the summed activity of these neurons aligned to stimulus onset (stimulus-triggered PETH) shows low-latency sharp peaks, with the peak within 20 ms of stimulus onset followed by a rapid delay; the response is broad and suppressed when aligned to movement onset, Fig. 4b-c (black traces). The tiny fraction of neurons identified as stimulus responders highlights that internally generated dynamics rather than externally driven responses massively dominate whole-brain activity. This finding is consistent with the small fraction of brain-wide activity variance explained by sensory responses via decoding analyses in [20], but shows in granular detail that the number of neurons involved is also small.

The next response pattern in the temporal arc of a trial is a band of activity following on from the stimulus response (Fig. 4a, blue box). This group is distinct from the stimulus responders even though the cells fire contiguously after the stimulus responders without a clear boundary between them in the stimulus presentation period, because they do not respond post-stimulus during the feedback period. They also belong to a distinct k-means functional cluster, (Fig. 3b), and unlike stimulus responders, the summed response of these neurons exhibits its largest peak when aligned to movement onset (blue curve, Fig. 4b-c), suggesting this group is linked to movement with higher temporal specificity than to the stimulus. These neurons, which we call integrators, exhibit a ramping response build-up over the stimulus period to a consistent level when movement-aligned (Fig. 4b-c, blue traces). Most significant for our assignment of these neurons as integrators is the termination of the ramp in firing rate just before movement onset, Fig. 4c: the responses ramp, peak, and then drop just as the movement is about to begin. The peak response occurs at a consistent and short latency before movement onset across trial types (Fig. 4b), fittingly for an integrator whose peak response determines the decision to move.

The chronologically third and fourth activity patterns we highlight in the brain's activity arc over a trial belong to what we will call movement initiation neurons (Fig. 4a, cyan box) and movement neurons (Fig. 4a, green box). The movement initiation cells are distinguished from the integrators by becoming active at a fixed latency before movement, peaking right at movement onset, then fading over the course of movement, Fig. 4b,c, cyan curves. This group may include motor preparatory neurons [27]. The response of these movement initiators contrasts with actual movement neurons, which are active only during movement, Fig. 4a (green box). Movement neurons only begin activity ramping around the integrator neuron peak, Fig. 4b-c. Their activity then continues to build strongly over the early movement period and persists during movement, Fig. 4c.

Next, motivated by the observation that input-driven changes account for only $\sim 1-2\%$ of the variance in neural activity, movement-related signals for $\sim 20\%$, and the remainder reflects unknown but coordinated internal dynamics [21, 28], we sought to understand the nature of this activity. We found a large and striking neuron group, consisting of neurons that fire in contextually specific activity sequences that begin at stimulus onset, and tile the interval from stimulus onset to movement onset, Fig. 4a (purple box). The cross-validated rastermap, in which neural orderings are determined from an entirely different set of trials than shown, is critical to establish that this is a real group because even random responses can be organized into some sequence. The presence of sequential activity patterns in the cross-validated rastermap means that a sequential activity pattern in one split of trials is present in a disjoint trial split, indicating that the relative activation order of neurons in the sequence is preserved across trials. Consistent with the preservation of sequential activity across trials splits, the same sequences are

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

present for different trial types (e.g. stimulus-aligned response columns for left or right stimulus and right or left choice), which necessarily consist of distinct trials Fig. 4a. The sequential activity patterns are contextually specific, in that there are four distinct sets of neurons with sequential activity tiling the stimulus period in the trial (cf. the vertically stacked set of four activity sequences in the purple block of Fig. 4a, such that on each of the four mistake trial types (choice side opposite the stimulus, labels colored gray in Fig. 4a), one of these sequences has higher activity (cf. movement-aligned trial intervals). Thus, we call this group contextual internal sequence/error-state neurons. These sequential responses are distinct from the combined spatiotemporally evolving responses of stimulus responders and integrator neurons, which also tile the stimulus period: the summed activity of stimulus responders and integrators combines a sharp transient and a slow ramp Fig. 4b,c (black and blue curves), while it is almost entirely flat for the sequence neurons, because they more-evenly tile the stimulus interval, Fig. 4b,c (purple curve). If the neural sequences are intrinsically generated timers rather than stimulus-driven and integrating responses, we should be able to better decode time from them than from the sensory and integrating responders, which indeed is true: the rapidly evolving activation of stimulus responders permits accurate temporal decoding once they have been activated, but only after their 20 ms onset latency, in contrast to the immediate decodability of time from the sequence neurons Fig.S9. Moreover, while the dense and widespread integrator responses enabled accurate decoding of time in the middle of the stimulus period, decoding accuracy dropped off at later times as the integrator saturated; the sequence neurons enabled more accurate decoding, with an accuracy that remained the same over the whole stimulus period. Next, the sequential neurons have consistently much lower firing rates than the stimulus responders and integrators, consistent with a separate identity, Fig.3f. Finally, the brain areas to which these internal sequences-error-state neurons belong are notably medial entorhinal cortex, amygdala, higher-order isocortical areas like insula, retrosplenial cortex, temporal association areas, and hippocampus, Fig.4d, Fig.S22.

Finally, we focus on neurons that predict a wrong outcome, before movement and error feedback. We find that cerebellum is relatively well-represented in regional participation of these error neurons, Fig. 4d (Fig.S21), and the cerebellar subfields participating most strongly in error prediction are distinct from the cerebellar subfields with strong representation of cells involved in movement initiation and integration.

In total, our rastermap-based analysis identifies $\sim 734/50240$ or 1.46% neurons as stimulus responders, $\sim (2860 + 1150)/50240$ or 7.98% neurons as movement initiators and movement neurons, and $\sim (2300 + 20650 + 4100)/50240$ or 53.84% neurons as integrator, contextual sequence processing and error-coding neurons over the course of a trial. In other words, internally generated brain dynamics that impose structure on the world during the trial dominate neural processing, with refined numbers from rastermap in line with previous estimates that motor activity dominates over sensory activity in explaining neural variance [21, 7, 12, 11, 29]. Looking collectively at which neurons participate in which functions, Fig. 4d, it is clear that for the IBL task, cortex is relatively unimportant. The integrator neurons are primarily subcortical rather than cortical, suggesting it would be fruitful to examine those areas in decision making studies, and that for simple evidence integration tasks, areas like PPC and LIP may be downstream readouts of subcortical decisions [30, 31]. Nevertheless, cortex and hippocampus apparently remain ready to help the brain adapt to new contexts and perform new learning by expressing contextual temporal scaffolds linked to task structure, Fig. 4d.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

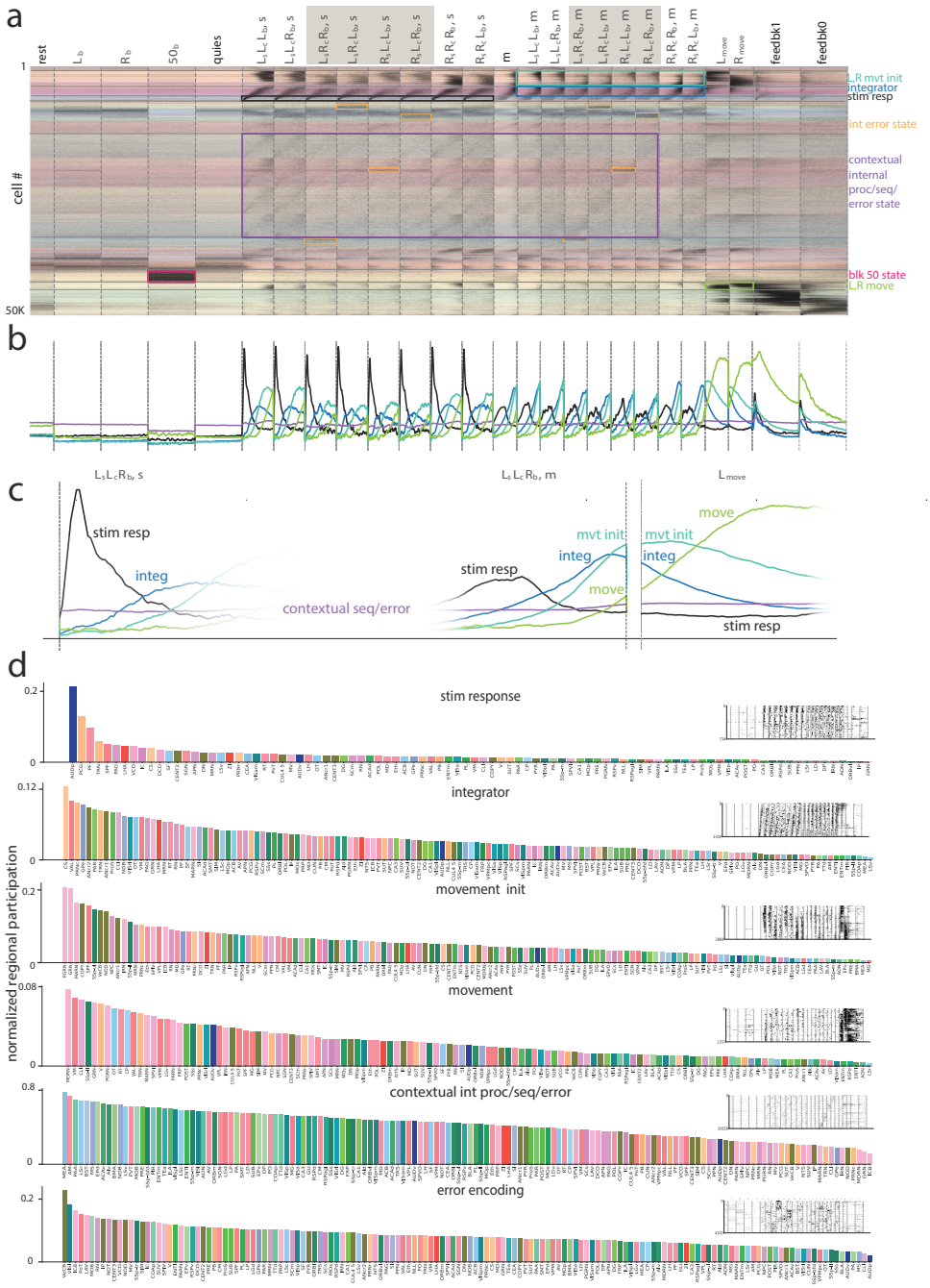


Figure 4. Finely resolved functional classes and regional participation. a Neural feature vectors ordered by rastermap [26] (cross-validated rastermap plot: sorting based on randomly selected 50% of the trials, with responses visualized for remaining 50%). All visible features of the rastermap from Fig.3 are preserved except for the single strong diagonal line. However, there are clear diagonal bands preserved across trials and trial types/contexts. Refined functional categories enabled by rastermap, shown as colored boxes with labels. **b** Averaged feature vectors for the refined functional categories, for stimulus response, integrator, movement initialization, movement, and internal processing/sequence/error groups. **c** Magnified stimulus-aligned and movement-aligned segments of the averaged feature vectors for stimulus response, integrator, movement initialization, and movement groups for two trial types. **d** Regional participation in each of the refined functional categories, measured as a normalized regional participation score (definition: fraction of cells in a region that belong to the functional category).

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Regional functional specialization differences

We have seen that cells belonging to a given functional category are dispersed across the brain, and in any brain region there are cells belonging to all functional response categories, questioning the idea of specialization of regions to specific functions in mice. We return to the question of regional specialization to ask whether there are systematic distinctions between brain regions, by considering distributional differences. For each of the ~ 200 Beryl regions, we considered the distribution of cell functional categories defined by the k-means functional clusters, Fig. 5a. Confirming the prior finding, all regions contained most functional categories but there was a diversity in the concentrations of distributions: some regions exhibit a notably broad, non-specialist character while others are relatively more specialized (e.g. region SIM in Fig. 5a has a nearly flat distribution, while MS has more than double the count in cluster 4 than any other cluster; distributions for all regions shown in Fig. S1. From these distributions, we can compute a specialization score (1 minus the entropy of its normalized distribution, which lies in the $[0, 1]$ interval for each region), with 1 corresponding to strong specialization (all cells of one functional type) and 0 to no specialization (equal distribution of all functional types). Not surprisingly, the distribution of specialization scores is peaked near 0, Fig. 5c.

When we overlay the specialization scores onto the Swanson map, side-by-side with a map of Cosmos levels, we see that there is not a clear relationship between Cosmos categories and functional specialization, Fig. 5d. Regions in the cerebellum (brown, e.g. SIM) and hindbrain (orange, e.g. PRNr) tend to have low specialization, while those with high specialization tend to be cortical (green, e.g. MOB), however specialization does not co-cluster with Cosmos level. To more quantitatively relate these results to the conventional view that lower brain regions are more functionally specialized while higher brain regions are generalists, we directly plot functional specialization scores against a structural hierarchy score obtained for a subset of brain regions via tract-tracing (Fig. 6d in [32]). Here, we find structural hierarchy and specialization are not significantly related, Fig. 5g.

To obtain verification of this result that functional specialization and structural hierarchy do not correlate, and show that it holds independent of our functional clustering technique, we computed a separate specialization score for each Beryl region based on a decoding analysis of 5 task variables (stimulus side, choice, feedback, wheel-speed, wheel-velocity; see Fig. S9 in [20] for details). Some regions had flatter distributions of decoding scores (e.g. region PRNr) than others (e.g. region PA), Fig. 5b (see Fig. S3 for all regions). We compute the decoding specialization score (1 minus the entropy of the normalized decoding accuracy distribution), and find that it strongly correlates ($p<0.01$) with the specialization score for the cluster counts, Fig. 5d, e: thus, both measures of regional specialization are consistent. The R^2 between the decoding-based specialization score and structural hierarchy is again low, 0.04, with no significant trend.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

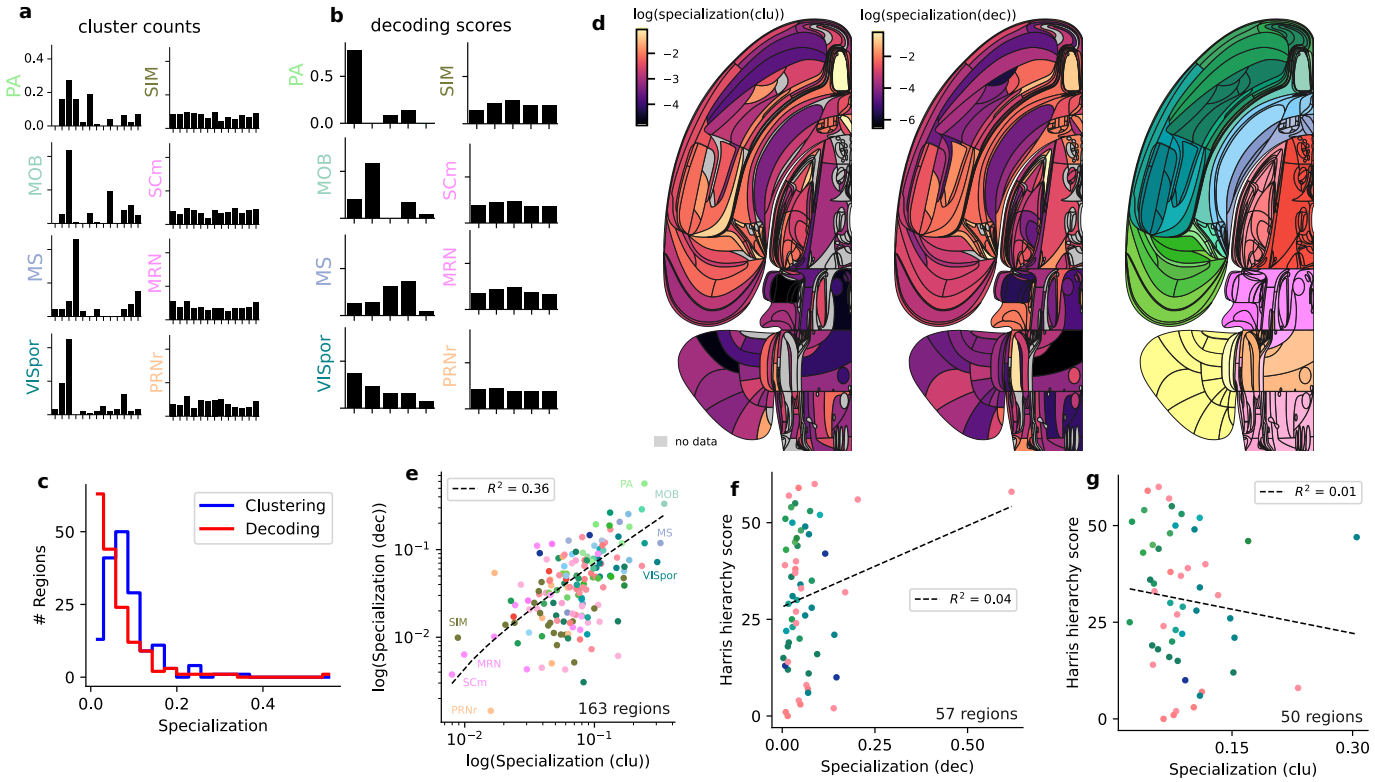


Figure 5. Differential amounts of region-level mixed selectivity in cortex a 13 k-means cluster cell counts (bars) for example regions; first (2nd) column regions with least (most) flat distributions, normalized to sum to 1. We define specialization as 1 minus the entropy of normalized distribution. **b** Trial variable decoding scores (from Fig. S9 in[20]) for example regions; first (2nd) column regions with least (most) flat distributions. Variables (bars): stimside, choice, feedback, wheel-speed, wheel-velocity, normalized accuracies to sum to 1. **c** Histograms of specialization scores across regions for cluster counts (clu) and decoding scores (dec), each peaking near zero. **d** Swanson flatmaps of all brain regions, colored by clu specialization (left), dec specialization (middle) and anatomical Cosmos color. Note the visible correlation of clu and dec specialization scores while no clear correlation with anatomical color blocks is visible. **e** Clu and dec specialization correlate clearly across 163 Beryl regions, supporting each as a measure of functional specialization. Regions with high specialization score tend to be cortical. **f**(g) Regional dec (clu) specialization versus structural hierarchy in cortical regions. No significant correlation with the structural hierarchy score (Fig. 6d in [32]) is found.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Dynamically evolving functional brain networks

We have seen that distinct computations unfold in different task phases and involve partially separate neuron groups. A natural question is how different parts of the brain coordinate during these distinct phases. The field of human fMRI, with a longer history of brain-wide recordings, has approached this question by performing multi-region functional graph analyses to identify functional subnetworks. The IBL dataset enables similar analyses, but at much higher ($\sim 10^3\times$ or ~ 1 ms versus ~ 1 s) temporal and spatial ($\sim 10^3\times$, with $\sim 60K$ cellular measurements in the supersession versus ~ 60 voxels) resolution, and with the inclusion of subcortical structures. We now seek an area-wise analysis of dynamical functional networks, or how the brain forms subnetworks to compute over different phases of the task, at high precision.

To assay functional network-level interactions, we defined a functional similarity score between (Beryl) regions per task phase as a proxy for the strength between them, focusing on the following conditions: the quiescent period; the stimulus response period in congruent trials (where the stimulus side equals the block side); the stimulus response period in surprise trials (when the stimulus side is opposite the block side, at least 10 trials deep within each block); during movement initialization; in the movement period; and during feedback after correct and wrong trials (see Tab.1 for detailed task phase definitions). We computed this similarity on region-wise cellular activity maps, which comprised smoothed 2-D embeddings of the Z-scored cellular feature vectors (as in Fig. 1d) per region and per task phase (Methods and Fig. 6a, Fig.S5). Similarity was computed as the cosine distance (defined in Methods) between maps, which yielded a similarity matrix across regions per task phase. As before, Z-scoring ensured that baseline firing rate differences were eliminated, allowing us to focus solely on functionally similar response patterns. We then applied Louvain community detection to the similarity matrices to identify functional networks (region clusters with similar responses), and sorted neurons to be adjacent if they were assigned to the same community, Fig. 6b-e (top-left in each panel). To assess which of the identified networks were reliable and meaningful, we applied three types of control analyses (see Methods), as follows. Only networks that are deemed significant by the intersection of the three controls below are numbered in Fig. 6b-e.

First, we identified which networks had a high 'network stability' score through split-trial comparisons (Figs. S13-S14). We repeated Louvain clustering across several trial splits; for each identified cluster (community) in each split, we defined a binary regional participation vector (whether the region was in the community), and used these to create a regional membership similarity matrix across all clusters and all splits for a condition. We assigned a network stability score ($\in [0, 1]$) defined by the average value of its regional membership similarity score (2), and only communities with a stability score exceeding a threshold were deemed statistically stable.

Second, to verify that the resulting networks identified as stable are different than would be obtained by random chance, we shuffled the region labels of all cells before computing the region-wise functional similarity scores and Louvain clustering. The resulting groupings were visually much more numerous and smaller (Fig. S13) than in the unshuffled data (Fig. S11- S12, right two columns), and their stability scores were well below the stability score threshold (3).

Our third control was based on the statistical significance of the functional responses of the identified networks relative to shuffles. We computed each network's PETH as a weighted (by region participation scores) average of the PETHs of the member regions, Fig. 6b-e (line plot in each panel), and compared these to PETHs generated by randomly sampling the same number of regions from all recorded Beryl regions in the brain Figs S11,S12. A network was deemed significant in its functional properties if its PETH was significantly different from these the random network PETHs for the majority of time bins.

We find that within each condition, brain states typically form a few non-orthogonal but distinct and regionally distributed subnetworks, suggesting distinct processing modes, Fig. 6b-e (networks viewed as undirected graphs, middle-left panels). The regional composition of identified subnetworks varies substantially from one task phase and condition to the next, meaning that the functional groupings are transient and rapidly shift across time and condition in the task, Fig. 6b-e (Swanson maps, right panels; darker colors indicate higher participation score). For instance, the quiescence period exhibits of a multiplicity of small, decoupled, marginally significant subnetworks, with the sole significant network being predominantly subcortical (hindbrain), Fig. 6b . During the congruent stimulus phase (Fig. 6c), the three dominant subnetworks consist of hippocampus–thalamus–dorsal striatum; cerebellum–midbrain–palladium-hindbrain; and neocortex-hypothalamus, while in the surprise stimulus phase, the networks collapse to two dominant ones, one remaining similar to the cortical-hypothalamic one and the other integrating

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

the hippocampal-dorsal striatal-thalamic network with the cerebellum-midbrain-palladium-hindbrain one. This suggests that during surprise, the hippocampus becomes engaged with subcortical processing, consistent with [33, 34].

Movement initiation and movement, adjacent task phases, involve quite different subnetworks, reflecting dynamic reconfiguration (Fig. 6d): the movement initiation period exhibits three distinct subnetworks, one primarily neocortical engaging early sensory processing and somatosensory areas, but also interacting with lateral hypothalamus, inferior colliculus, cerebellum, and medulla structures like GRN; the second is predominantly hippocampal-thalamic-dorsal striatum (CP) with somatosensory-secondary motor cortex engaged with hindbrain; and the third is predominantly cerebellar with engagement of superior colliculus. Brain states in the movement period collapse into two statistically robust subnetworks, one largely similar to the first movement initiation subnetwork, and the other a mixture of hippocampus, high-level association areas (RSP), medulla (PAG), and hypothalamus. Finally, from comparing the correct and incorrect feedback phase subnetworks (Fig. 6e), we find that despite the large differences in inputs, cellular response tuning and network PETHs that separate these conditions, they share an identical subnetwork (networks 3, 1 respectively) that is similar to the third movement network, presumably reflecting movement-related computations leaking through from the movement phase immediately earlier. In what remains, there is significant but not perfect similarity between the subnetworks 2 in both feedback conditions; these exhibit high brainstem-cerebellar-striatal activation and possibly show common aspects of feedback processing. Subnetwork 1 in the correct feedback phase is distinct, involving cortico-cerebellar activation with widespread neocortical activity across low and high level cortical areas.

One consistent pattern is that the within-condition subnetworks in every condition recapitulate the green-pink or cortical-subcortical response divide that we saw earlier at the single-cell level. This is evident in the graphs of Fig. 6, as well as when we form a similarity matrix of the region membership of subnetworks from all these and other task conditions, Fig. 6f. The subnetwork regional similarity matrix has a two-block structure, each with a representative subnetwork from most block conditions. The upper block corresponds to primarily subcortical processing networks, and the lower one to primarily cortical processing networks. Interesting, we see that congruent stimulus network 1 overlaps strongly with prestimulus prior network 1 and correct feedback network 3, consistent with the possibility that a congruent stimulus and a correct response build up the strength of the block prior, as predicted from a neural circuit model of the IBL task [35].

Finally, we compare these functional brain networks to established fMRI-based networks from the mouse brain [32, 36, 37, 38, 39, 40, 41, 10, 42, 43, 44] to determine whether there is a correspondence, Fig. 6g. The fMRI sensory mode network (smn, Fig. 6g, third row) reassuringly overlaps with a subnetwork from each of the surprise and congruent stimulus conditions, but also with the more-cortical subnetworks from each of the movement initiation and movement conditions. However, beyond this similarity, we report few meaningful relationships between fMRI-defined subnetworks and our high spatio-temporal resolution brain networks, even when comparing e.g. our quiescent state network with the fMRI default mode networks (Fig. 6g, bottom row), which correspond to similar brain states. The quantified dominance of subcortical processing in the brain throughout the course of the IBL task, the sparsity of fMRI maps in the mapping to Beryl regions even for cortical structures, and the relative lack of subcortical recordings in the fMRI studies are the likely causes of the discrepancy.

Together these results show, in a spatiotemporally precise and brain-wide way that all brain areas participate in all functions, and that the brain forms emergent collective dynamical coalitions that rapidly reconfigure between task phases to enable accurate perceptual decision-making.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

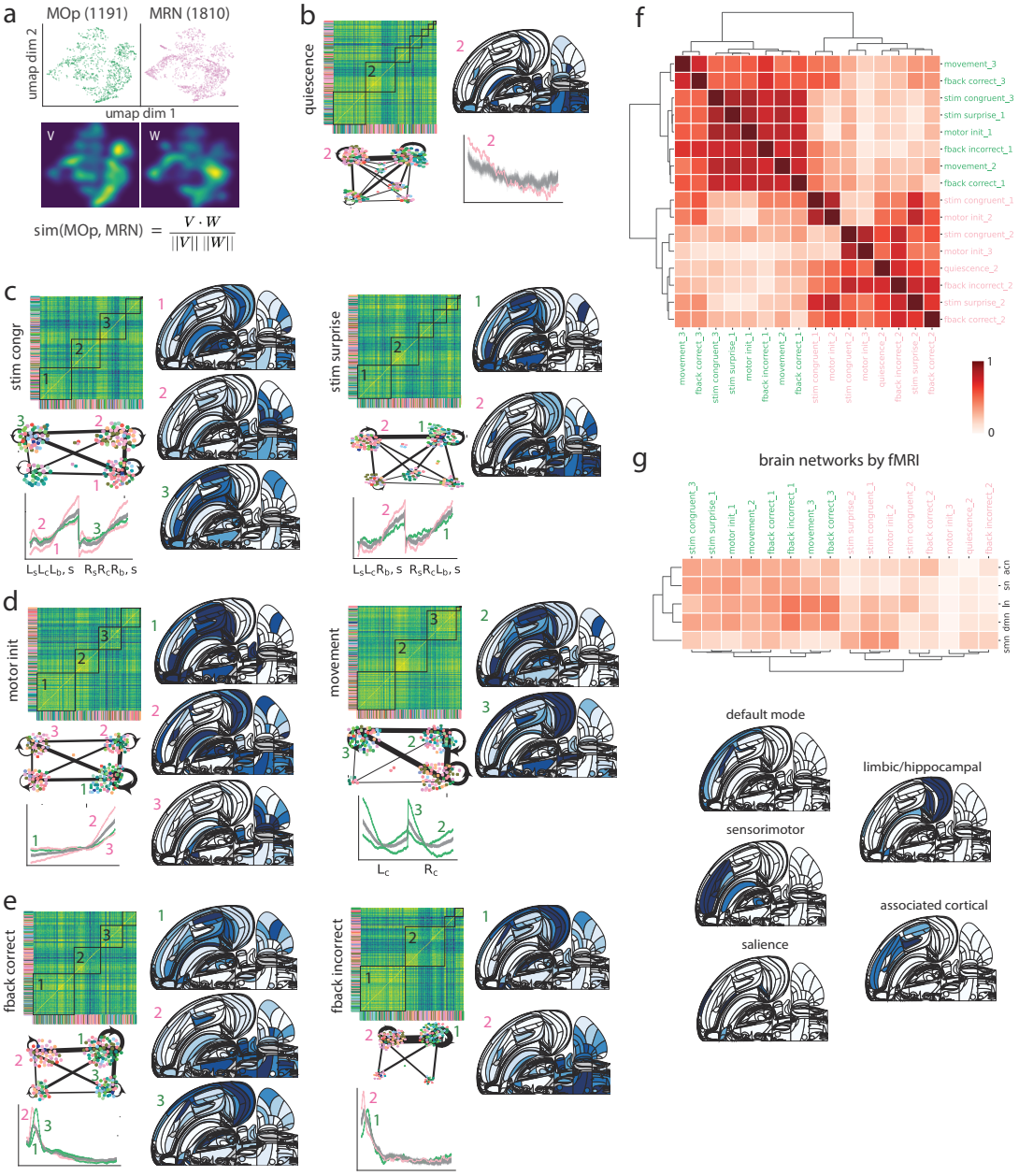


Figure 6. Dynamically reconfiguring networks across different phases of the task. a To compare functional responses between pairs of regions, we calculate a cosine similarity score of two regions' cellular responses from a smoothed 2D embedding for the relevant task phase (here, shown for the concat condition which includes all task phases; see Fig.S5 for more examples). **b-e** The regional similarity matrices are ordered by the Louvain community detection algorithm to identify groups of brain regions (networks) for various subsets of PETHs (task phases) shown in panels (b) through (e) (see Tab.1 for task phase definitions). The algorithm consistently shows a cortical-subcortical separation in two (or three) major networks across phases of the task, as evident in the network graphs (some clusters mostly pink-subcortical, and the others mostly green-cortical + hippocampus) and Swanson maps showing regional membership in these clusters (darker color indicates stronger membership from split-trial cross-validation). However the exact network membership varies across different task phases/ conditions, reflecting dynamic reconfiguration of the brain's networks. We also show line plots of the average PETHs for each network as compared to shuffled controls (gray), and the networks are only included in the current figure if they pass all three rounds of control analyses as described in the Methods. **f** Summary table showing cosine similarity of Swanson maps between pairs of our defined networks, having two clear groups that correspond to the cortical-subcortical separation. **g** Cosine similarity of Swanson maps between our defined networks and canonical fMRI networks in the mouse brain (Swanson maps below), showing low similarity throughout. Colors in the Swanson maps indicate how consistently we find the regions being identified in a network in the literature, with darker color being more consistent. [36, 37, 38, 39, 41, 10, 45, 42, 43, 44].

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Discussion

We have shown that individual neurons fall into distinct and interpretable clusters based on their functional responses during the IBL task. These categories largely involve one aspect of the task: responding to the stimulus, integration or choice generation, movement initiation, movement generation, predicted error coding, and representation of internal biases based on prior expectation. Ordering neurons by the similarity of their responses, via the rastermap, afforded a glimpse at cellular resolution of an interpretable view of the brain-wide flow of computational processing. Distinct populations of neurons transiently respond to and process one aspect of the task, then pass the ball to another set, from storing expectations to perception, to decision-making, then action. We found that these transient processing steps were widely distributed across the brain. Thus, most individual neurons do not appear to exhibit mixed selectivity, or at least are dominated by selectivity to single sub-functions within the task.

Neurons with similar functional responses are not spatially localized: they are dispersed across the brain, and each brain region, defined by fine (Beryl) categories, contains all functional classes of neurons. Thus, there is a bidirectionally low co-correspondence between function and anatomy. Each region comprises neurons with diverse responses, and thus presumably participates in most computations. Interestingly therefore, though there are clear regional variations in task-independent electrophysiological features, which relate to cell types and gene expression [25], these do not translate into corresponding differences in functional roles as determined by functional responses. Note however, that the observation of disperse functional responses does not establish that these responses are generated in a distributed way: it remains possible that there are more localized generators of various functional responses, and these generators broadcast the signal for downstream use [13].

The primary regional distinction we find in functional responses is between cortical and subcortical processing, with integration, choice, and motor initiation dominated by subcortical processing. Thus, our results hint that cortical processing may be relatively unimportant for expert performance on the IBL task.

Our data directs us away from a region-dominant view in which a given anatomically or cytoarchitecturally defined region specializes in its own set of computations, and after performing these computations on its inputs it passes these outputs to the next region. The findings bolster instead the view of brainwide processing in which each computation involves neurons across a large number of areas, and each area participates in a large number of functions. Thus, our data encourages us to reconceptualize computation in the mouse brain as forming distinct processing pathways or loops through the brain, each of which winds through and involves many or most brain areas within each phase of a task, as in Fig.1f, consistent with recent observations of e.g. large cortical-cerebellar loops forming a combined dynamical system for cognitive processing and control [46] and as hypothesized elsewhere [47].

With the advent of several large, high resolution brain-wide data-sets across species and the computational ability to build neural “foundation” models, e.g. [48], we expect future work to support our claim of brain-wide distributed networks, showing above-chance representation and decodability of behavior and relevant task variables from neural activity in most anatomical brain structures.

There is nevertheless some finer regional structure. First, at the distributional level: some regions have flatter distributions of all functional neurons, while others have a less uniform distribution, pointing toward greater specialization. However, we found no relationship between previous anatomical measures of cortical hierarchy and these distributional specializations. Second, at the level of co-activation in different task phases: examining co-active regions in each task phase, in the style of fMRI ’network’ analyses [10], we found that there were indeed distinct patterns of co-activation or correlation between regions at different task phases, a set of emergent collective computational modes. Our data permitted us to also include subcortical participation in these networks, and to parse the evolution of correlated patterns at high spatial and temporal resolution. We found reliable patterns of co-active networks in our across-trial, across-subject analyses with rigorous controls, and these reinforced the dominant cortical-subcortical distinction that we have seen throughout this work. We saw that these networks evolve or reconfigure rapidly within a trial as different computations unfold. The resulting networks exhibit some but relatively weak overlap with fMRI-derived brain networks, for reasons that require future investigation. There are numerous methodological differences resulting in distinct data resolutions, from levels of spatiotemporal precision, to differences in the detected activity signal (blood flow versus neural spiking) to different amounts of coverage of the brain (cortical versus brain-wide). The resulting

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

set of networks form detailed atlases for future studies on brain-wide coordination for different functions. In addition, it will be interesting to perform more-detailed comparisons across species, given the spatial localization of function observed in humans, for instance in face [4] and language [49] encoding.

Finally, we have discovered that the bulk of neural activation in the brain during the task is not related to immediate externally driven sensory responses or the more widespread encoding of movement generation [21, 28], but rather consists of internally generated sweeping scalar signals across the brain. The regions participating in this internal signal suggests that these sequences might be broadcast from the hippocampal complex, which seems possible given reports that hippocampus generates sequential activity starting from any salient task event, in a context-specific way [50, 51]. Given that all brain regions participate in these temporally structured and contextually specific sweeping signals, it implies that all regions exhibit neurons with mixed selectivity and contextual sensitivity, in line with theoretical models of functional specialization of network [52].

These findings reveal a distributed and dynamic computational architecture in which the brain orchestrates task-relevant processing through shifting, functionally defined networks, dominated by internal states, that transcend anatomical boundaries.

Methods and Materials

Data and task

We analyzed the brain wide map dataset of the international brain laboratory (IBL), consisting of recordings from 139 mice in 12 labs performing a decision-making task with sensory, motor, and cognitive components, obtained with 699 Neuropixels probe insertions covering 279 brain areas in the left forebrain and midbrain and the right hindbrain and cerebellum. The dataset is publicly available (download instructions; data browsing online tool at https://viz.internationalbrainlab.org) with full details here[20]. Find also information about histologically obtained anatomical coordinates there.

The International Brain Laboratory (IBL) decision-making task[19] is such that on each trial, a visual stimulus appeared on the left or right on a screen, and the mouse had to move it to the center by turning a wheel with its front paws within 60 seconds. The prior probability for the stimulus to appear on the left/right side was constant over a block of trials, at 20/80% (right block) or 80/20% (left block). Blocks lasted for between 20 and 100 trials, drawn from a truncated geometric distribution (mean 51 trials). Block changes were not cued. Stimulus contrast was sampled uniformly from five possible values (100, 25, 12.5, 6.25, and 0%). The 0% contrast trials, when no stimulus was presented, were assigned to a left or right side following the probability distribution of the block, allowing mice to perform above chance by incorporating this prior in their choices. Following a wheel turn, mice received positive feedback in the form of a water reward, or negative feedback in the form of a white noise pulse and a $2\,\mathrm{s}$ time-out. The next trial began after a delay, followed by a quiescence period during which the mice had to hold the wheel still.

Peri-event time histograms (PETHs)

We binned spiking activity of neurons and trials declared as "good" as in[20] with time bins of length 12.5 ms, and stride 2 ms. See data loading functions `load_good_units` for neural data and `load_trials_and_mask` for trial data in `bwm_loading.py`. Firing rates were squared prior to averaging across trials, to pronounce heterogeneity. We decided to choose a set of 26 PETH types (tab1) such that a balance is struck between maximizing trial-specific information while keeping as many sessions as possible. I.e. too specific PETH criteria are satisfied by too few sessions. Our choice resulted in 569 unique insertions being included in the analysis and grouped together, each neuron being described by a functional feature vector of all 26 PETHs concatenated together, then z-scored to homogenize differences in baseline activity. We used 61473 neurons in the single cell analyses and 163 regions, when restricting to regions with at least 50 neurons, in the regional analyses.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Table 1. PETH definitions Each PETH is defined by an alignment event, the trials to average, and the window. The two times in the window are relative to the alignment event, with the left value being positive when before the event and the right value being positive if after the event. "Deep block" indicates averaged trials must be at least 10 trials after a block change. "Phase" indicates subsets of PETHs.

Name	Alignment Event	Trial Types to Average	Window (s)	Phase
rest	stimOn_times	All trials	-1.15 to 1	Resting
L_b	stimOn_times	ProbabilityLeft = 0.8 (block left)	-0.4 to 0.1	Pre-stim prior
R_b	stimOn_times	ProbabilityLeft = 0.2 (block right)	-0.4 to 0.1	Pre-stim prior
50_b	stimOn_times	ProbabilityLeft = 0.5	-0.4 to 0.1	Block 50
quies	stimOn_times	All trials	-0.4 to 0.1	Quiescence
$L_sL_cL_b$, s	stimOn_times	Stim left, choice left, block left, deep block	0 to 0.2	Stim congruent
$L_sL_cR_b$, s	stimOn_times	Stim left, choice left, block right, deep block	0 to 0.2	Stim surprise
$L_sR_cR_b$, s	stimOn_times	Stim left, choice right, block right, deep block	0 to 0.2	Mistake
$L_sR_cL_b$, s	stimOn_times	Stim left, choice right, block left, deep block	0 to 0.2	Mistake
$R_sL_cL_b$, s	stimOn_times	Stim right, choice left, block left, deep block	0 to 0.2	Mistake
$R_sL_cR_b$, s	stimOn_times	Stim right, choice left, block right, deep block	0 to 0.2	Mistake
$R_sR_cR_b$, s	stimOn_times	Stim right, choice right, block right, deep block	0 to 0.2	Stim congruent
$R_sR_cL_b$, s	stimOn_times	Stim right, choice right, block left, deep block	0 to 0.2	Stim surprise
m	firstMovement_times	All trials	-0.15 to 0	Motor initiation
$L_sL_cL_b$, m	firstMovement_times	Stim left, choice left, block left	-0.15 to 0	
$L_sL_cR_b$, m	firstMovement_times	Stim left, choice left, block right	-0.15 to 0	
$L_sR_cR_b$, m	firstMovement_times	Stim left, choice right, block right	-0.15 to 0	Mistake
$L_sR_cL_b$, m	firstMovement_times	Stim left, choice right, block left	-0.15 to 0	Mistake
$R_sL_cL_b$, m	firstMovement_times	Stim right, choice left, block left	-0.15 to 0	Mistake
$R_sL_cR_b$, m	firstMovement_times	Stim right, choice left, block right	-0.15 to 0	Mistake
$R_sR_cR_b$, m	firstMovement_times	Stim right, choice right, block right	-0.15 to 0	
$R_sR_cL_b$, m	firstMovement_times	Stim right, choice right, block left	-0.15 to 0	
L_{move}	firstMovement_times	Choice left	0 to 0.15	Movement
R_{move}	firstMovement_times	Choice right	0 to 0.15	Movement
feedbk1	feedback_times	Positive feedback	0 to 0.3	Feedback correct
feedbk0	feedback_times	Negative feedback	0 to 0.3	Feedback incorrect

Umap and rastermap embeddings

We embedded the 60 k cells' feature vectors (concatenated PETHs, 2496 time bins in total) in 2 dimensions using `umap.UMAP(n_neighbors=15, n_components=2)`[53] implemented in python.

The rastermap algorithm[26] was used in python to sort the 60 k feature vectors by similarity, using these parameters: `Rastermap(n_PCs=200, n_clusters=100, locality=0.75, time_lag_window=5, bin_size=1)`. We visualized the sorted vectors using `matplotlib.imshow(data, vmin=0, vmax=1.5, cmap="gray_r", aspect="auto")`.

Functional distance metric for brain region pairs

We measure the functional distance between brain regions as follows. For a given region, the 2D umap embedding of all cells in this region (across recordings) is rasterized into a 256 x 256 pixel grid, with the cell density per pixel, then smoothed by applying a 2D Gaussian kernel with standard deviation of 10 pixels in each dimension and normalized such that the maximum grayscale per pixel is set to 1 and the minimum to zero. Given two such images for two regions, the images are flattened into vectors $\mathbf{v_1}, \mathbf{v_2}$ and the cosine similarity of the two vectors is defined as the functional distance between regions:

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

$$d_{1,2} = \frac{\mathbf{v_1 v_2}}{|\mathbf{v_1}||\mathbf{v_2}|}$$

Louvain clustering

We use `sknetwork.clustering.Louvain` implemented in python to identify regional networks in the functional similarity matrices of different trial phases. We use the following parameters in the presented analysis: `resolution=1.01, random_state=0`, after sweeping through parameters and other clustering algorithms that result in less reproducible networks with lower adjusted random index (`sklearn.metrics.adjusted_rand_score`) and adjusted mutual information (`sklearn.metrics.adjusted_mutual_i`

Statistical significance of networks

To assess which of the identified networks were reliable and meaningful, we applied three types of control analyses.

First, we identified which networks had a high 'network stability' score through split-trial comparisons, as follows. We split all trials of a given type into two groups to obtain two independent feature vectors for each cell. Within each split, we re-computed similarity matrices and Louvain clusters (communities). We repeated this process four times to generate a total of eight sets of leave-half-out communities for each condition (each cross-validation set based on a random subset of $1/2$ the total trials). We compared region participation across the eight sets of communities to define a network stability score, as follows (see Methods for details): All of the communities (for a given condition) in all the eight splits were assigned binary region participation vectors (of length equaling the number of Beryl regions), with 1 indicating a region's membership in the community and 0 otherwise, and we used these to create a regional membership similarity matrix across all clusters in all splits for the condition (Figs. S13-S14), sorting the matrix elements by similarity to display the communities (SI Figs. S13-S14, column 1). We only look at communities with sizes > 6 when sorting the similarity matrix, meaning these communities are identified in at least 6 out of the 8 leave-half-out trial splits analysis. The network stability score of each community is defined as the average value (of the regional membership similarity score) within the community (2): conceptually, if the same areas always appear together in one community across splits, the stability score would be 1; otherwise, it would take some value in the interval $[0, 1]$. Communities with a stability score < 0.47 (the average stability score across communities and across trial phases) were deemed statistically unstable and not numbered or plotted as Swanson maps in the main or supplemental figures.

Second, to verify that the resulting stable identified networks are different than would be obtained by random chance, we shuffled the region labels of all cells before computing the region-wise functional similarity scores and Louvain clustering. The resulting groupings were visually much more numerous and smaller (Fig. S13) than in the unshuffled data (Fig. S11- S12, right two columns). We quantitatively compared the shuffled data networks with the identified networks by computing the stability scores for the shuffled data networks. All of the shuffled networks have stability scores well below the criteria of stability score > 0.47 (3).

Our third control was based on the statistical significance of the functional responses of the identified networks. The goal is to ensure the networks are a result of distinct and related responses of brain regions within a community rather than community-independent regional response variations. For each community, we computed its cluster PETH by a weighted average of the PETHs of the member regions (weighted by the participation scores of the regions, as defined in the next section in Methods), Fig. 6b-e (line plot in each panel). We compared the community PETH with 1000 control PETHs generated by randomly sampling the same number of regions as in the community from across the set of all recorded Beryl regions in the brain, and computing their average PETHs weighted by the same set of regional participation scores (randomly assigned to the random set of regions), Figs S11,S12. At each time bin, two p-values are defined by the fraction of times the shuffled controls are larger or smaller than the true PETH, and the true PETH is significantly different from the controls if either of the p-values are smaller than 0.01. A community was deemed significant only if the resulting network PETHs were significantly different from the shuffled controls for the majority of the analysis time window.

Only networks that are deemed significant by the intersection of the three conditions above are numbered in Fig. 6 of the main text. In Swanson maps, we show the membership of regions in each significant network, weighted by a network participation score.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Regional participation score of networks

We define a region's participation score within a given network as the fraction of times the region is identified within the network across the eight-fold split-trial controls. If a region is identified in a given network in all eight splits, its participation score in the network is 1, and if identified in none, its score is 0. For each network that passed the statistical significance tests, we show the Swanson maps of them in the main results (fig.6) by coloring the regions with the regional participation score.

Supervised decoding analysis

We decoded 5 types of classes (restricted to label classes with at least 20 neurons) from either the concatenated PETHs or 30 concatenated trial-independent ephys features. The trial-independent ephys features were channel-specific, i.e. several neurons may sit on the same channel, and consisted of the following (see full definitions in cite ephysTF manuscript??):

```
'alpha_mean',  'alpha_std',      'depolarisation_slope', 'peak_time_secs',
 'peak_val',    'polarity',       'psd_alpha',           'psd_alpha_csd',
 'psd_beta',    'psd_beta_csd',   'psd_delta',           'psd_delta_csd',
 'psd_gamma',   'psd_gamma_csd',  'psd_lfp',             'psd_lfp_csd',
 'psd_theta',   'psd_theta_csd',  'recovery_slope',      'recovery_time_secs',
 'repolarisation_slope', 'rms_ap', 'rms_lf',             'rms_lf_csd',
 'spike_count_x', 'spike_count_y', 'tip_time_secs',      'tip_val',
 'trough_time_secs', 'trough_val'
```

The classes were:

Beryl Anatomical region parcellation into 204 regions (see [20] for full list)
Cosmos Anatomical region parcellation into 10 coarse regions (Fig. 2g and full definition in
Canonical 7 groups of Beryl regions by similarity (from Fig. 1 in [24])

```
        'Prefrontal': ['FRP', 'ACAd', 'ACAv', 'PL', 'ILA', 'ORBl', 'ORBm', 'ORBvl'],
        'Lateral': ['AId', 'AIv', 'AIp', 'GU', 'VISc', 'TEa', 'PERI', 'ECT'],
        'Somatomotor': ['SSs', 'SSp-bfd', 'SSp-tr', 'SSp-ll',
        'SSp-ul', 'SSp-un', 'SSp-n', 'SSp-m', 'MOp', 'MOs'],
        'Visual': ['VISal', 'VISl', 'VISp', 'VISpl', 'VISli', 'VISpor', 'VISrl'],
        'Medial': ['VISa', 'VISam', 'VISpm', 'RSPagl', 'RSPd', 'RSPv'],
        'Auditory': ['AUDd', 'AUDp', 'AUDpo', 'AUDv']
        'Other': [all other regions]
```

K-means 13 functional k-means clusters as shown in Fig. 1e
Layers Partition into 6 cortical layers (defined by number in Beryl region acronym: 1: 2k, 2, 1k, 3: 3 k, 4: 1k, 5: 6k, 6: 0.1 k cells), thalamus (8 k cells) and rest (40 k cells)

We used logistic regression (`sklearn.linear_model.LogisticRegression` L2 regularization, C=1.0, scikit-learn) with 5-fold stratified cross-validation, scaling data within each fold using scikit-learn StandardScaler. Classifier performance was evaluated using mean accuracy across folds. Shuffled-label controls were included to assess classification significance. 10 decoding runs were sufficient, given the low variance across folds.

Acknowledgments

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

[1] Enny H van Beest, Sreedeep Mukherjee, Lisa Kirchberger, Ulf H Schnabel, Chris van der Togt, Rob RM Teeuwen, Areg Barsegyan, Arne F Meyer, Jasper Poort, Pieter R Roelfsema, et al. Mouse visual cortex contains a region of enhanced spatial resolution. *Nature communications*, 12(1):4029, 2021.

[2] Carl CH Petersen. The functional organization of the barrel cortex. *Neuron*, 56(2):339–355, 2007.

[3] Troy A Hackett, Tania Rinaldi Barkat, Barbara MJ O'Brien, Takao K Hensch, and Daniel B Polley. Linking topography to tonotopy in the mouse auditory thalamocortical circuit. *Journal of Neuroscience*, 31(8):2983–2995, 2011.

[4] Nancy Kanwisher, Josh McDermott, and Marvin M Chun. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of neuroscience*, 17(11):4302–4311, 1997.

[5] Doris Y Tsao, Winrich A Freiwald, Roger BH Tootell, and Margaret S Livingstone. A cortical region consisting entirely of face-selective cells. *Science*, 311(5761):670–674, 2006.

[6] Reza Rajimehr, Haoran Xu, Asa Farahani, Simon Kornblith, John Duncan, and Robert Desimone. Functional architecture of cerebral cortex during naturalistic movie watching. *Neuron*, 112(24):4130–4146, 2024.

[7] Ana R Inácio, Ka Chun Lam, Yuan Zhao, Francisco Pereira, Charles R Gerfen, and Soohyun Lee. Brain-wide presynaptic networks of functionally distinct cortical neurons. *Nature*, pages 1–11, 2025.

[8] Dheeraj S Roy, Young-Gyun Park, Minyoung E Kim, Ying Zhang, Sachie K Ogawa, Nicholas DiNapoli, Xinyi Gu, Jae H Cho, Heejin Choi, Lee Kamentsky, et al. Brain-wide mapping reveals that engrams for a single memory are distributed across multiple brain regions. *Nature communications*, 13(1):1799, 2022.

[9] James M Stafford, Benjamin R Jarrett, Oscar Miranda-Dominguez, Brian D Mills, Nicholas Cain, Stefan Mihalas, Garet P Lahvis, K Matthew Lattal, Suzanne H Mitchell, Stephen V David, et al. Large-scale topology and the default mode network in the mouse connectome. *Proceedings of the National Academy of Sciences*, 111(52):18745–18750, 2014.

[10] Jennifer D Whitesell, Adam Liska, Ludovico Coletta, Karla E Hirokawa, Phillip Bohn, Ali Williford, Peter A Groblewski, Nile Graddis, Leonard Kuan, Joseph E Knox, et al. Regional, layer, and cell-type-specific connectivity of the mouse default mode network. *Neuron*, 109(3):545–559, 2021.

[11] Susu Chen, Yi Liu, Ziyue Aiden Wang, Jennifer Colonell, Liu D Liu, Han Hou, Nai-Wen Tien, Tim Wang, Timothy Harris, Shaul Druckmann, et al. Brain-wide neural activity underlying memory-guided movement. *Cell*, 187(3):676–691, 2024.

[12] Andrei Khilkevich, Michael Lohse, Ryan Low, Ivana Orsolic, Tadej Bozic, Paige Windmill, and Thomas D Mrsic-Flogel. Brain-wide dynamics linking sensation to action during decision-making. *Nature*, 634(8035):890–900, 2024.

[13] Camden J MacDowell, Alexandra Libby, Caroline I Jahn, Sina Tafazoli, Adel Ardalan, and Timothy J Buschman. Multiplexed subspaces route neural activity across brain-wide networks. *Nature communications*, 16(1):3359, 2025.

[14] Somayeh Shahsavarani, David N Thibodeaux, Weihao Xu, Sharon H Kim, Fatema Lodgher, Chinwendu Nwokeabia, Morgan Cambareri, Alexis J Yagielski, Hanzhi T Zhao, Daniel A Handwerker, et al. Cortex-wide neural dynamics predict behavioral states and provide a neural basis for resting-state dynamic functional connectivity. *Cell reports*, 42(6), 2023.

[15] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106, 1962.

[16] Jia Zhu, Hana Hasanbegović, Liu D Liu, Zhenyu Gao, and Nuo Li. Activity map of a cortico-cerebellar loop underlying motor planning. *Nature neuroscience*, 26(11):1916–1928, 2023.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[17] Mattia Rigotti, Omri Barak, Melissa R Warden, Xiao-Jing Wang, Nathaniel D Daw, Earl K Miller, and Stefano Fusi. The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585–590, 2013.

[18] Valerio Mante, David Sussillo, Krishna V Shenoy, and William T Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474):78–84, 2013.

[19] The International Brain Laboratory, Valeria Aguillon-Rodriguez, Dora Angelaki, Hannah Bayer, Niccolo Bonacchi, Matteo Carandini, Fanny Cazettes, Gaelle Chapuis, Anne K Churchland, Yang Dan, et al. Standardized and reproducible measurement of decision-making in mice. *Elife*, 10, 2021.

[20] International Brain Laboratory, Brandon Benson, Julius Benson, Daniel Birman, Niccolo Bonacchi, Kcénia Bougrova, Sebastian A Bruijns, Matteo Carandini, Joana A Catarino, Gaelle A Chapuis, et al. A brain-wide map of neural activity during complex behaviour. *biorxiv*, pages 2023–07, 2023.

[21] Simon Musall, Matthew T Kaufman, Ashley L Juavinett, Steven Gluf, and Anne K Churchland. Single-trial neural dynamics are dominated by richly varied movements. *Nature neuroscience*, 22(10):1677–1686, 2019.

[22] David Raposo, Matthew T Kaufman, and Anne K Churchland. A category-free neural population supports evolving demands during decision-making. *Nature neuroscience*, 17(12):1784–1792, 2014.

[23] Charles Findling, Felix Hubert, International Brain Laboratory, Luigi Acerbi, Brandon Benson, Julius Benson, Daniel Birman, Niccolò Bonacchi, Sebastian Bruijns, Matteo Carandini, et al. Brain-wide representations of prior information in mouse decision-making. *BioRxiv*, pages 2023–07, 2023.

[24] Pierre Le Merre, Sofie Ährlund-Richter, and Marie Carlén. The mouse prefrontal cortex: Unity in diversity. *Neuron*, 109(12): 1925–1944, 2021.

[25] The International Brain Laboratory. paper-ephys-atlas. https://github.com/int-brain-lab/paper-ephys-atlas, 2025. Accessed: 2025-06-16.

[26] Carsen Stringer, Lin Zhong, Atika Syeda, Fengtong Du, Maria Kesa, and Marius Pachitariu. Rastermap: a discovery method for neural population recordings. *Nature Neuroscience*, pages 1–12, 2024.

[27] Mark M Churchland, M Yu Byron, Stephen I Ryu, Gopal Santhanam, and Krishna V Shenoy. Neural variability in premotor cortex provides a signature of motor preparation. *Journal of Neuroscience*, 26(14):3697–3712, 2006.

[28] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893, 2019.

[29] Nicholas A Steinmetz, Peter Zatka-Haas, Matteo Carandini, and Kenneth D Harris. Distributed coding of choice, action and engagement across the mouse brain. *Nature*, 576(7786):266–273, 2019.

[30] Joshua I Gold and Michael N Shadlen. The neural basis of decision making. *Annu. Rev. Neurosci.*, 30(1):535–574, 2007.

[31] Kenneth W Latimer, Jacob L Yates, Miriam LR Meister, Alexander C Huk, and Jonathan W Pillow. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science*, 349(6244):184–187, 2015.

[32] Julie A Harris, Stefan Mihalas, Karla E Hirokawa, Jennifer D Whitesell, Hannah Choi, Amy Bernard, Phillip Bohn, Shiella Caldejon, Linzy Casal, Andrew Cho, et al. Hierarchical organization of cortical and thalamic connectivity. *Nature*, 575(7781): 195–202, 2019.

[33] Talia Brandman, Rafael Malach, and Erez Simony. Default mode network dynamics track the cognitive state of surprise during naturalistic audiovisual stimulation. *Journal of Vision*, 20(11):348–348, 2020.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[34] Darya Frank, Alex Kafkas, and Daniela Montaldi. Experiencing surprise: The temporal dynamics of its impact on memory. *Journal of Neuroscience*, 42(33):6435–6444, 2022.

[35] Rylan Schaeffer, Mikail Khona, Leenoy Meshulam, Brain Laboratory International, and Ila Fiete. Reverse-engineering recurrent neural network solutions to a hierarchical inference task for mice. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 4584–4596. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/30f0641c041f03d94e95a76b9d8bd58f-Paper.pdf.

[36] Tomokazu Tsurugizawa, Kota Tamada, Nobukazu Ono, Sachise Karakawa, Yuko Kodama, Clément Debacker, Junichi Hata, Hideyuki Okano, Akihiko Kitamura, Andrew Zalesky, and Toru Takumi. Awake functional mri detects neural circuit dysfunction in a mouse model of autism. *Science Advances*, 6:eaav4520, 2020. doi: 10.1126/sciadv.aav4520.

[37] Valerio Zerbi, Joanes Grandjean, Markus Rudin, and Nicole Wenderoth. Mapping the mouse brain with rs-fmri: An optimized pipeline for functional network identification. *NeuroImage*, 123:11–21, 2015. doi: 10.1016/j.neuroimage.2015.07.090.

[38] Joanes Grandjean, Carola Canella, Cynthia Anckaerts, Gülebru Ayrancı, Salma Bougacha, Thomas Bienert, David Buehlmann, Ludovico Coletta, Daniel Gallino, Natalia Gass, Clément M Garin, Nachiket Abhay Nadkarni, Neele S Hübner, Meltem Karatas, Yuji Komaki, Silke Kreitz, Francesca Mandino, Anna E Mechling, Chika Sato, Katja Sauer, Disha Shah, Sandra Strobelt, Norio Takata, Tong Wu, Noriaki Yahata, Ling Yun Yeow, Yohan Yee, Ichio Aoki, M Mallar Chakravarty, Wei-Tang Chang, Marc Dhenain, Dominik von Elverfeldt, Laura-Adela Harsan, Andreas Hess, Tianzi Jiang, Georgios A Keliris, Jason P Lerch, Andreas Meyer-Lindenberg, Hideyuki Okano, Markus Rudin, Alexander Sartorius, Annemie Van der Linden, Marleen Verhoye, Wolfgang Weber-Fahr, Nicole Wenderoth, Valerio Zerbi, and Alessandro Gozzi. Common functional networks in the mouse brain revealed by multi-centre resting-state fmri analysis. *Neuroimage*, 205:116278, 2020. doi: 10.1016/j.neuroimage.2019.116278.

[39] Daniel Gutierrez-Barragan, Neha Atulkumar Singh, Filomena Grazia Alvino, Ludovico Coletta, Federico Rocchi, Elizabeth De Guzman, Alberto Galbusera, Mauro Uboldi, Stefano Panzeri, and Alessandro Gozzi. Unique spatiotemporal fmri dynamics in the awake mouse brain. *Current Biology*, 32:631–644.e6, 2022. doi: 10.1016/j.cub.2021.12.015.

[40] David Hike, Xiaochen Liu, Zeping Xie, Bei Zhang, Sangcheon Choi, Xiaoqing Alice Zhou, Andy Liu, Alyssa Murstein, Yuanyuan Jiang, Anna Devor, et al. High-resolution awake mouse fmri at 14 tesla. *Elife*, 13:RP95528, 2025.

[41] Hadi Vafaii, Francesca Mandino, Gabriel Desrosiers-Grégoire, David O'Connor, Xilin Shen, Ge Xinxin, Peter Herman, Fahmeed Hyder, Xenophon Papademetris, Mallar Chakravarty, Michael C Crair, R Todd Constable, Evelyn MR Lake, and Luiz Pessoa. Multimodal measures of spontaneous brain activity reveal both common and divergent patterns of cortical functional organization. *Nature Communications*, 15:229, 2024. doi: 10.1038/s41467-023-44363-z.

[42] Francesco Sforazzini, Adam J Schwarz, Alberto Galbusera, Angelo Bifone, and Alessandro Gozzi. Distributed bold and cbv-weighted resting-state networks in the mouse brain. *NeuroImage*, 87:403–415, 2014. doi: 10.1016/j.neuroimage.2013.09.050.

[43] Qasim Bukhari, Aileen Schroeter, David M Cole, and Markus Rudin. Resting state fmri in mice reveals anesthesia specific signatures of brain functional networks and their interactions. *Frontiers in Neural Circuits*, 11:5, 2017. doi: 10.3389/fncir.2017.00005.

[44] Joanes Grandjean, Valerio Zerbi, Joshua Henk Balsters, Nicole Wenderoth, and Markus Rudin. Structural basis of large-scale functional connectivity in the mouse. *Journal of Neuroscience*, 37:8092–8101, 2017. doi: 10.1523/JNEUROSCI.0438-17.2017.

[45] Ander Egimendia, Anuka Minassian, Michael Diedenhofen, Dirk Wiedermann, Pedro Ramos-Cabrer, and Mathias Hoehn. Aging reduces the functional brain networks strength—a resting state fmri study of healthy mouse brain. *Frontiers in Aging Neuroscience*, 11:277, 2019. doi: 10.3389/fnagi.2019.00277.

[46] Mark J Wagner and Liqun Luo. Neocortex–cerebellum circuits for cognitive processing. *Trends in neurosciences*, 43(1):42–54, 2020.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[47] Luiz Pessoa. *The entangled brain: How perception, cognition, and emotion are woven together*. MIT Press, 2022.

[48] Eric Y Wang, Paul G Fahey, Zhuokun Ding, Stelios Papadopoulos, Kayla Ponder, Marissa A Weis, Andersen Chang, Taliah Muhammad, Saumil Patel, Zhiwei Ding, et al. Foundation model of neural activity predicts response to new stimulus types. *Nature*, 640(8058):470–477, 2025.

[49] Alexander G Huth, Wendy A De Heer, Thomas L Griffiths, Frédéric E Theunissen, and Jack L Gallant. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600):453–458, 2016.

[50] Christopher J MacDonald, Kyle Q Lepage, Uri T Eden, and Howard Eichenbaum. Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, 71(4):737–749, 2011.

[51] Karyna Mishchanchuk, Gabrielle Gregoriou, Albert Qü, Alizée Kastler, Quentin JM Huys, Linda Wilbrecht, and Andrew F MacAskill. Hidden state inference requires abstract contextual representations in the ventral hippocampus. *Science*, 386 (6724):926–932, 2024.

[52] Gabriel Béna and Dan FM Goodman. Dynamics of specialization in neural modules under resource constraints. *Nature Communications*, 16(1):187, 2025.

[53] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

Supplementary Material

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



Figure S1. Neuron counts across K-means clusters for each brain region. For each brain region (Beryl parcellation) with more than 50 neurons, a histogram across 13 k-means clusters is shown, normalized such that all cluster counts add to 1. Note that nearly all regions contain neurons from all clusters.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

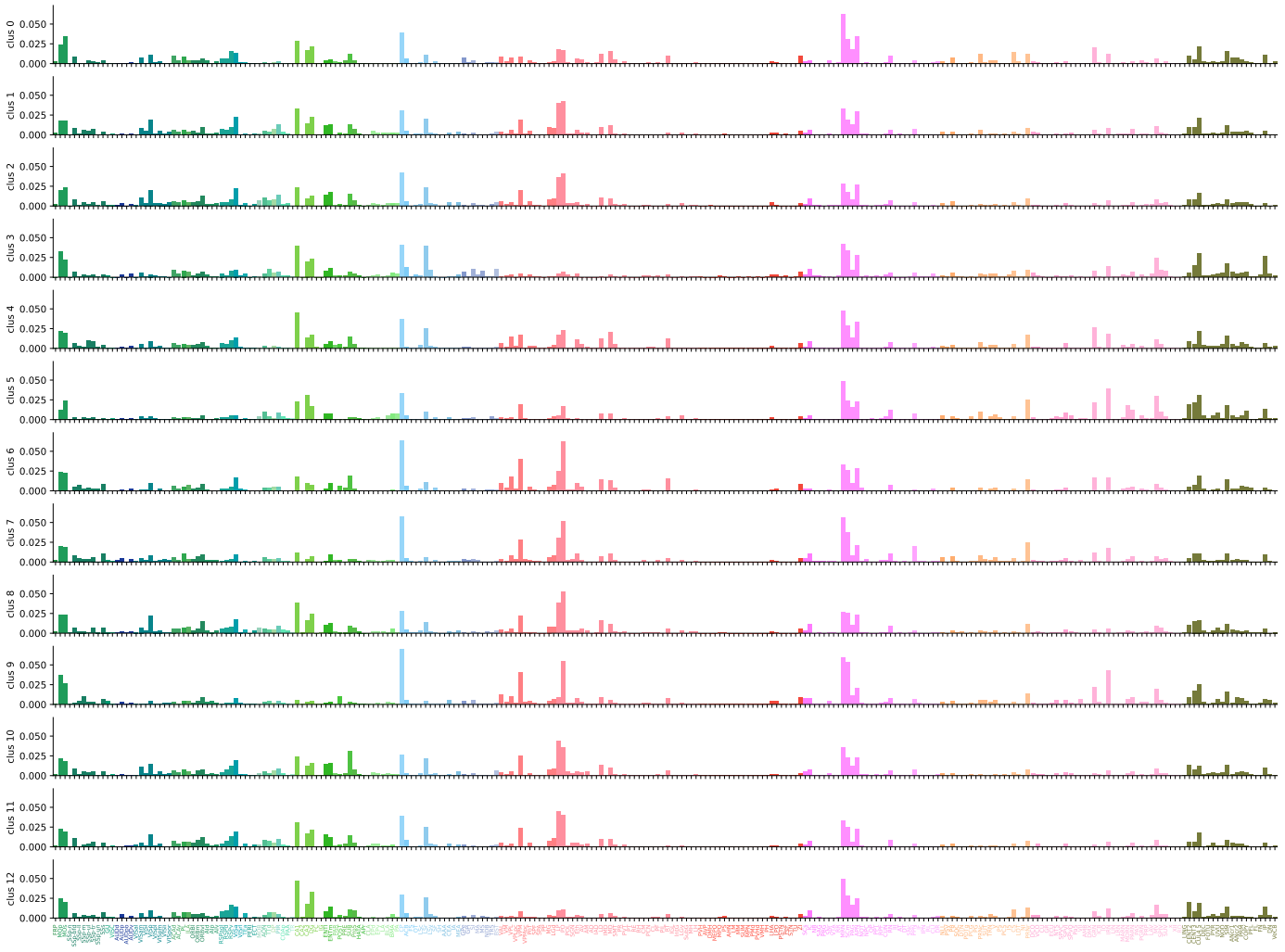


Figure S2. Neuron counts across brain region for each k-means cluster. The counts are normalized to sum to 1. Note the similarities in the shape of the distributions across the 13 clusters, supporting that neurons from each k-means cluster can be found in most Beryl regions.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

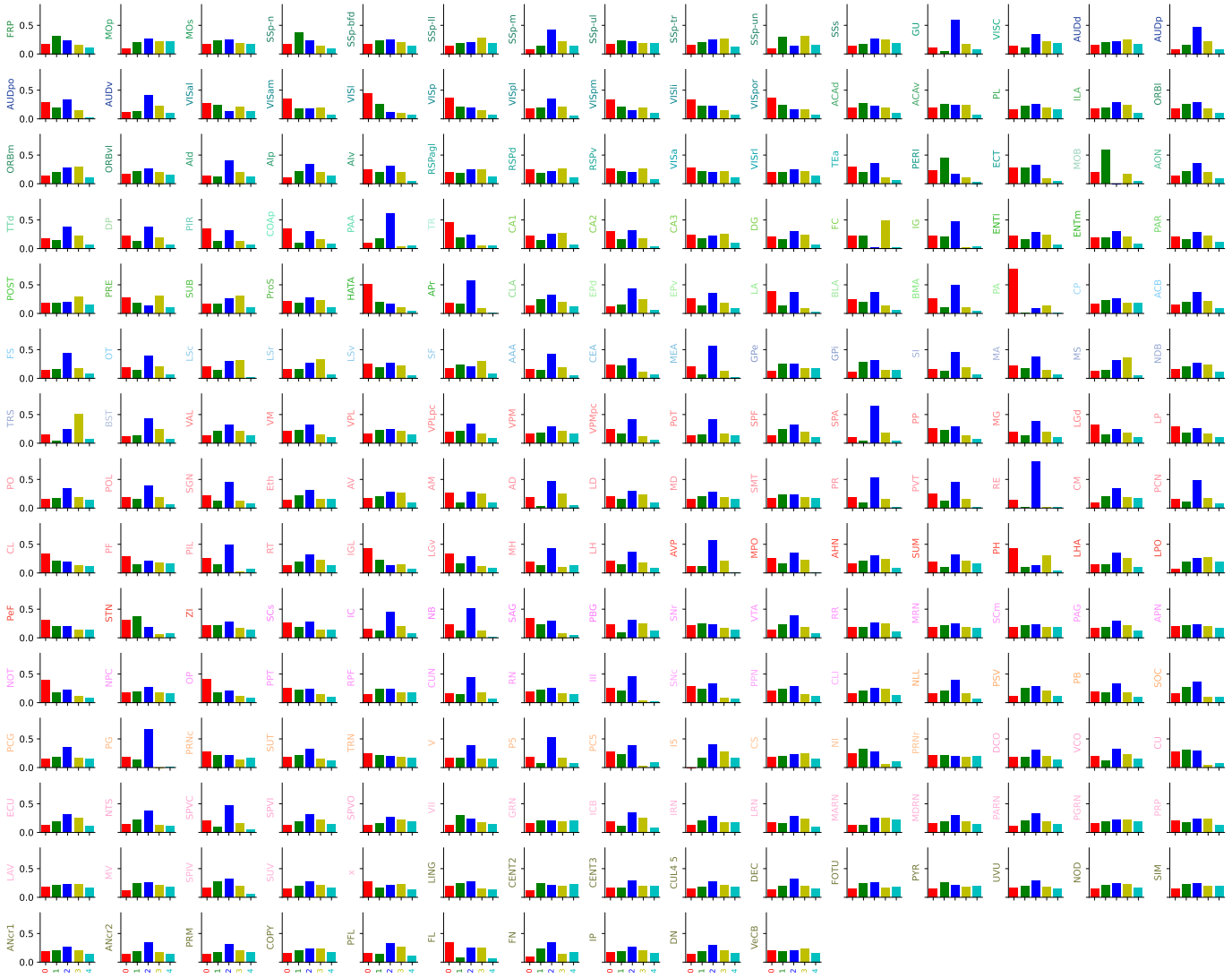


Figure S3. Decoding accuracies across task variables for each brain region. For each brain region (Beryl parcellation) with more than 50 neurons, the decoding score across 5 task variables is shown, being the average across all recordings of that region and normalized such that all decoding scores add to 1. The decoded task variables are: stimulus side, choice, feedback wheel speed, wheel velocity. Note that nearly all regions allow decoding of all task variables. See fig S9 in [20] for the per session results.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

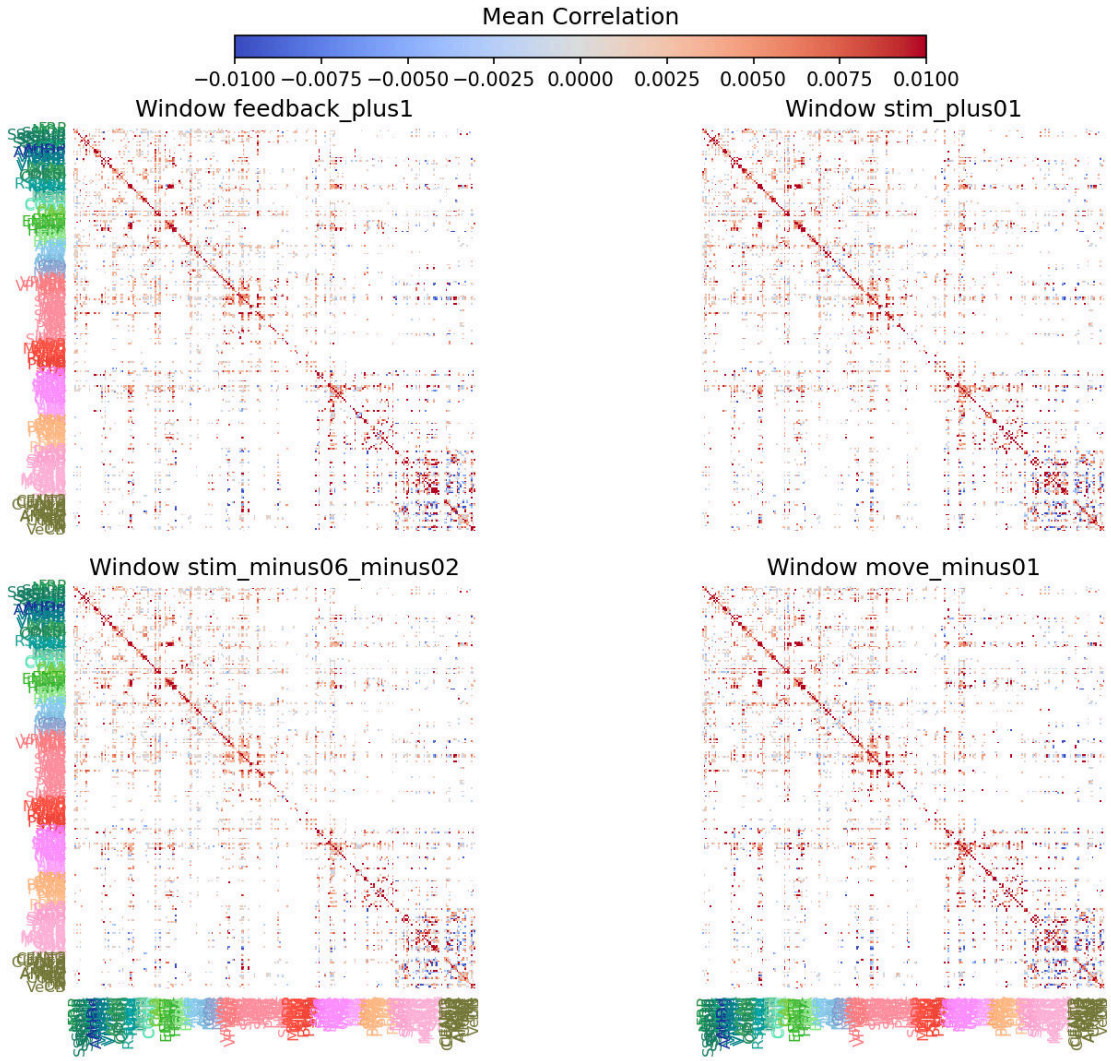


Figure S4. Binned time series correlation scores for Beryl region pairs barely differ across different task windows Pearson correlation score of binned times series of cell pairs, averaged across recordings and per Beryl region pairs, for a given window, say until 1 sec after feedback ("feedback_plus1"). The overall low correlation scores barely differ across the 4 task windows. Canonical coloring and ordering of Beryl region acronyms as in fig2 of [20]

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

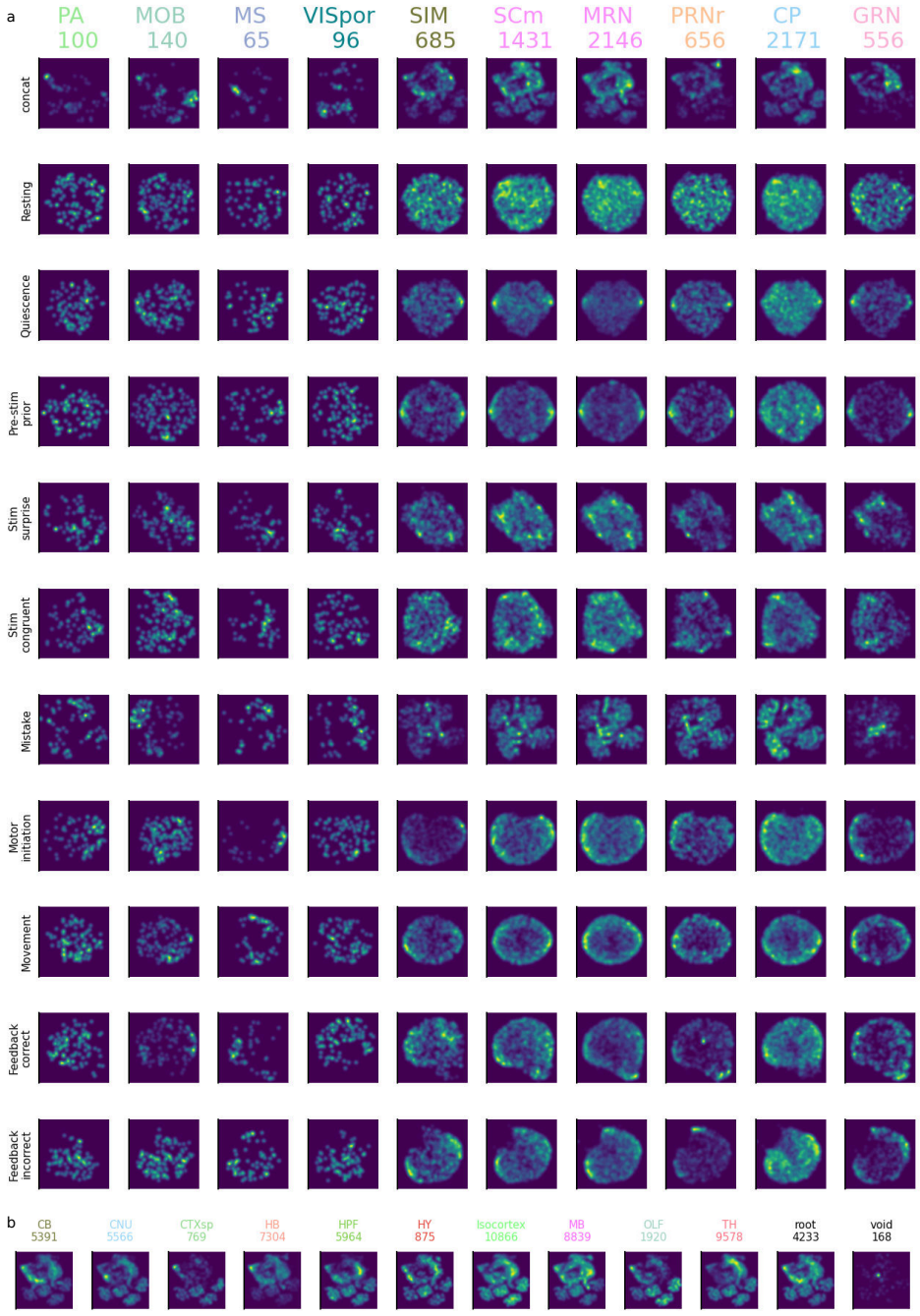


Figure S5. Smoothed 2d embeddings for example regions and networks a For 10 example regions and 11 networks (subsets of all PETHs, see Tab.1) the smoothed 2D umap embedding of the cell's feature vectors are shown, normalized such that each image's maximal pixel is 1 (see methods). The cosine similarities of these flattened images are used as a distance metric between regions in Fig.6. b For all concatenated PETHs (network is "concat") and all Cosmos regions, the smoothed and normalized 2D umap embedding of the cells is shown. Note the homogeneous distribution of isocortex and hippocampal cells (green title color, as cells in Fig2b) in contrast to midbrain (MB) and hindbrain (HB) cells, that show concentrations on top of embedding space. This functional/anatomical correlation is seen also as a green/pink color divide in Fig2b.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

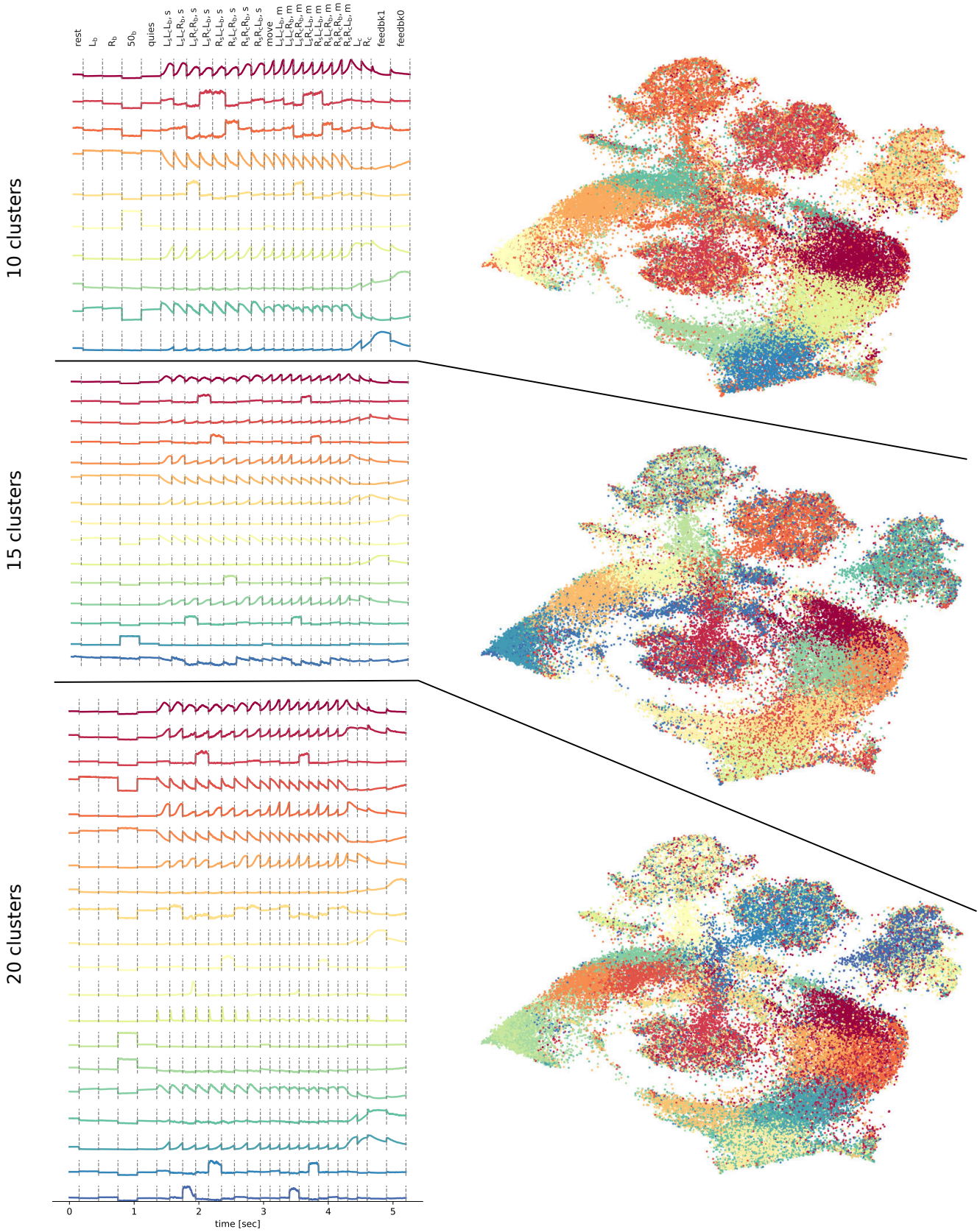


Figure S6. Varying the number of k-means clusters results in qualitatively similar functional groupings results. For 10, 15 and 20 k-means clusters, average feature vectors in the left column, colored umap 2D embeddings in right column. Note that 10 clusters group together mistake conditions while 20 subdivide them - see "islands" in umap landscape.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

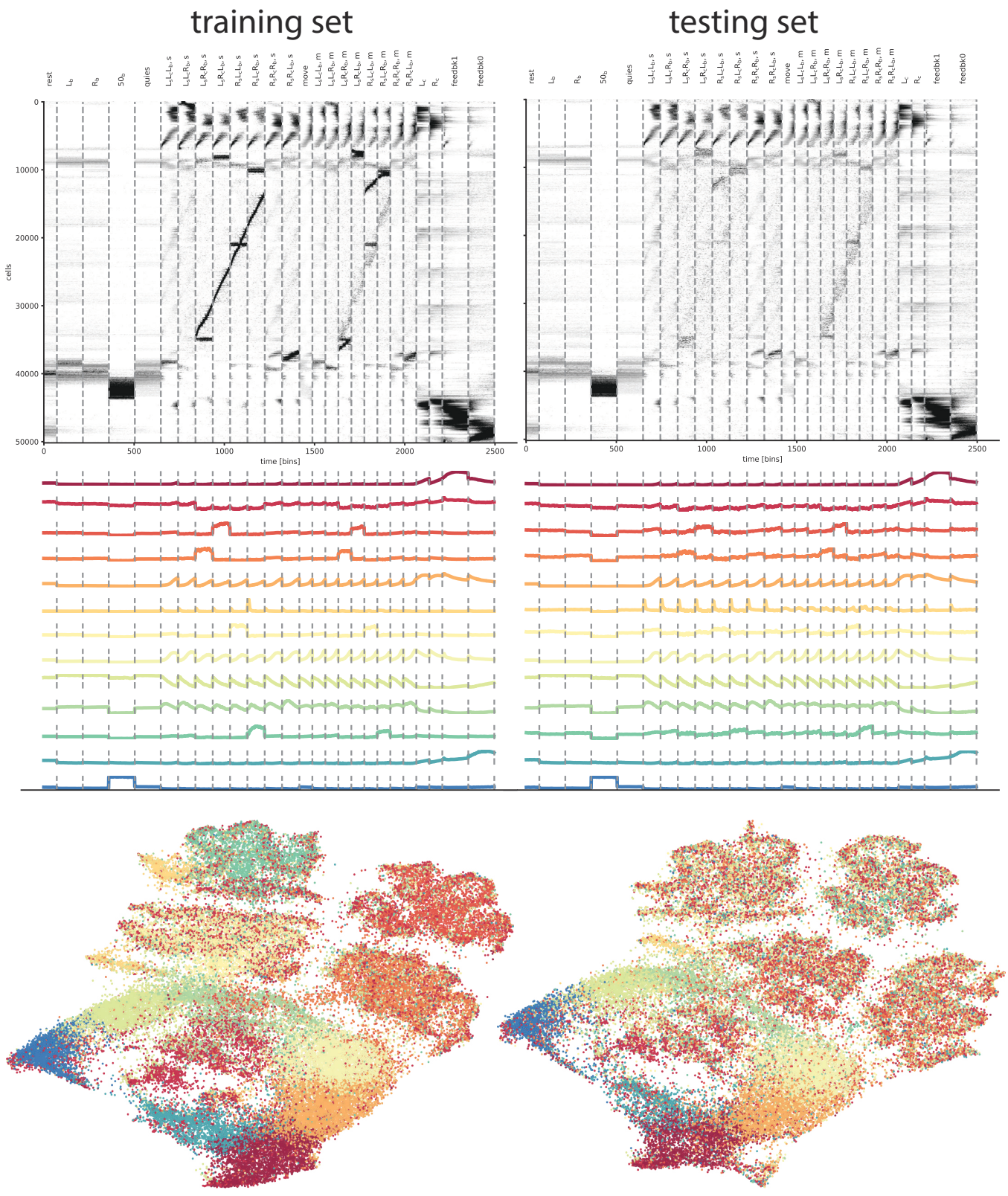


Figure S7. Split-trial cross validation of Rastermap and k-means Trials are randomly split into two sets (training and testing) for each alignment and trial type, from which two concatenated feature vectors are computed for each cell. Rastermap and k-means algorithms are applied on the training set, and we apply the same ordering and clustering labels to both the training and testing set. The majority of the patterns from Rastermap and clusters from k-means are preserved after cross-validation.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

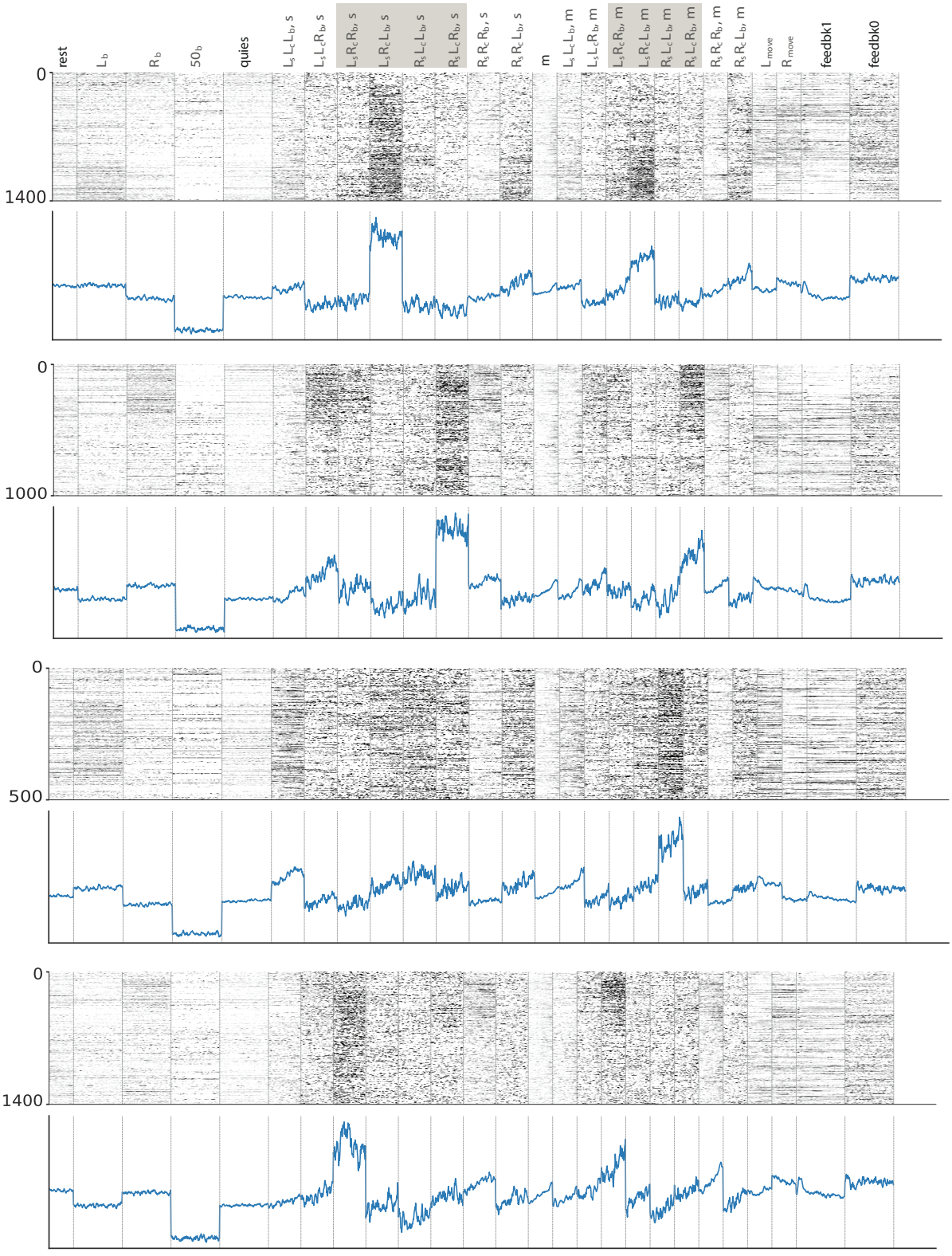


Figure S8. Error encoding cells for four mistake trial types We zoom in from fig.4d (error encoding panel) and show the raster and average PETH of four types of error encoding cells separately, based on the mistake trial types they respond to.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

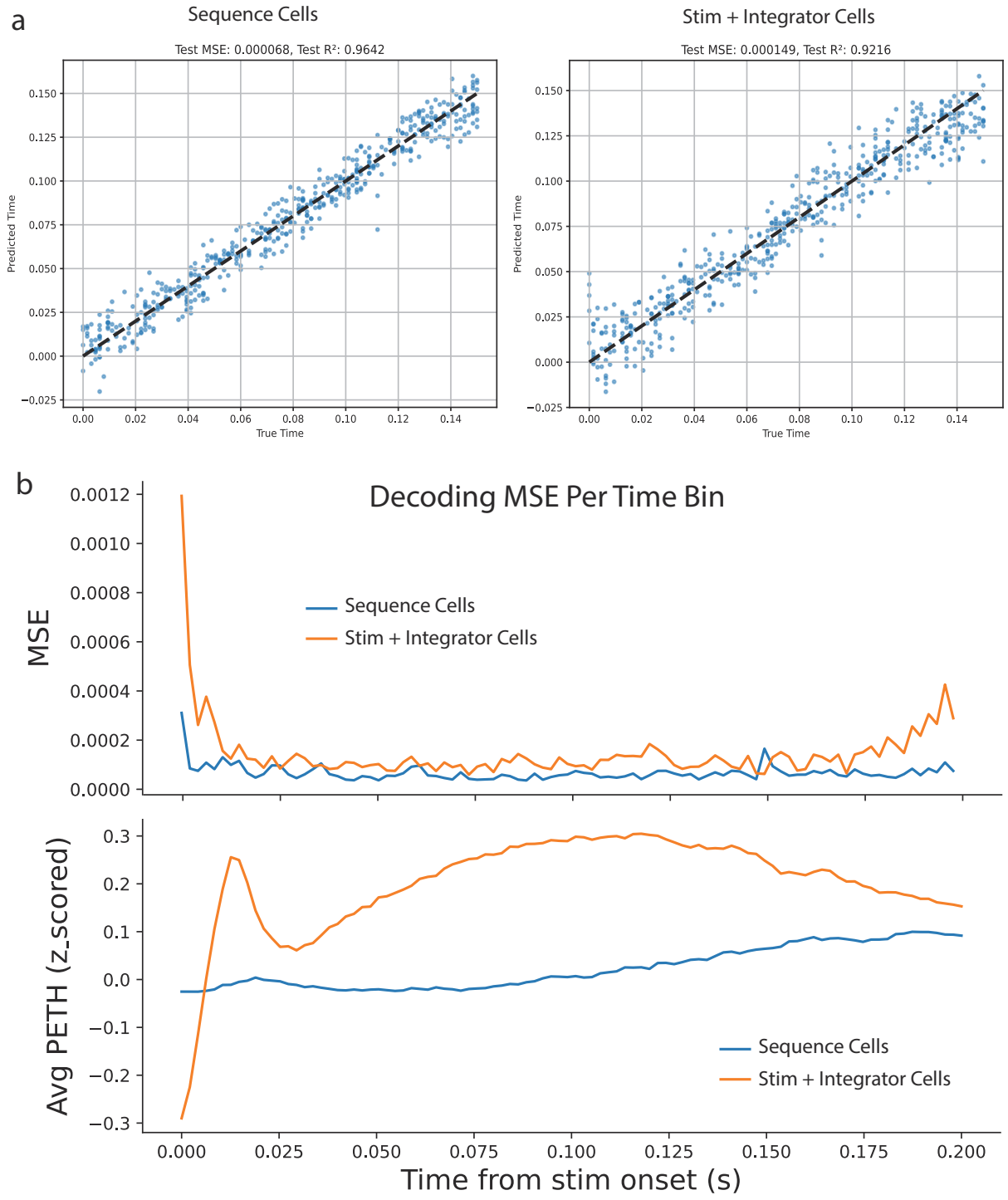


Figure S9. Decoding time from stimulus onset in sequence cells vs. stim response + integrator cells a. Both resulted in good performance with $R^2 > 0.9$. b. The mean-squared error is slightly higher for stim response + integrator cells compared to the sequence cells near the beginning and end of stimulus-onset period. This MSE difference marks the distinction between the stim response/ integrator cells from the sequence cells, and can be explained by differences in c. the average PETHs of these groups of cells: the stim response cells has a slight delay after stimulus onset to reach its peak response, and the integrator cells' activities start to decay near the end of stimulus period, whereas the sequence cells have a relatively stable level of activity after stimulus onset. Definitions of stim response, integrator, and sequence cells can be found in the main results section fig.4.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cluster	Stability
concat cluster 0	0.862
concat cluster 1	0.712
concat cluster 2	0.686
concat cluster 3	0.337
concat cluster 4	0.061
resting cluster 0	0.214
resting cluster 1	0.201
resting cluster 2	0.186
resting cluster 3	0.205
resting cluster 4	0.166
resting cluster 5	0.21
resting cluster 6	0.203
quiescence cluster 0	0.618
quiescence cluster 1	0.259
quiescence cluster 2	0.357
quiescence cluster 3	0.173
quiescence cluster 4	0.154
quiescence cluster 5	0.138
pre-stim prior cluster 0	0.903
pre-stim prior cluster 1	0.887
pre-stim prior cluster 2	0.205
pre-stim prior cluster 3	0.2
pre-stim prior cluster 4	0.204
mistake cluster 0	0.79
mistake cluster 1	0.632
mistake cluster 2	0.093
mistake cluster 3	0.694

Cluster	Stability
stim surprise cluster 0	0.541
stim surprise cluster 1	0.609
stim surprise cluster 2	0.402
stim surprise cluster 3	0.084
stim surprise cluster 4	0.269
stim surprise cluster 5	0.251
stim congruent cluster 0	0.784
stim congruent cluster 1	0.49
stim congruent cluster 2	0.503
stim congruent cluster 3	0.428
stim congruent cluster 4	0.069
motor init cluster 0	0.885
motor init cluster 1	0.719
motor init cluster 2	0.76
motor init cluster 3	0.423
movement cluster 0	0.849
movement cluster 1	0.734
movement cluster 2	0.675
movement cluster 3	0.061
fback correct cluster 0	0.601
fback correct cluster 1	0.616
fback correct cluster 2	0.599
fback incorrect cluster 0	0.708
fback incorrect cluster 1	0.636
fback incorrect cluster 2	0.437
fback incorrect cluster 3	0.368
fback incorrect cluster 4	0.037

Table 2. Stability scores for regional clusters. Stability scores identified from comparing clusters that are identified from different trial splits for all trial phases. Clusters with stability scores > 0.48 (marked as red) pass the stability criteria.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cluster	Stability
concat cluster 0	0.268
concat cluster 1	0.162
concat cluster 2	0.178
concat cluster 3	0.228
concat cluster 4	0.208
concat cluster 5	0.219
concat cluster 6	0.179
resting cluster 0	0.123
resting cluster 1	0.2
resting cluster 2	0.137
resting cluster 3	0.179
resting cluster 4	0.146
resting cluster 5	0.204
resting cluster 6	0.134
resting cluster 7	0.209
resting cluster 8	0.171
quiescence cluster 0	0.19
quiescence cluster 1	0.199
quiescence cluster 2	0.167
quiescence cluster 3	0.213
quiescence cluster 4	0.169
quiescence cluster 5	0.185
pre-stim prior cluster 0	0.206
pre-stim prior cluster 1	0.21
pre-stim prior cluster 2	0.15
pre-stim prior cluster 3	0.245
mistake cluster 0	0.18
mistake cluster 1	0.16
mistake cluster 2	0.179
mistake cluster 3	0.181
mistake cluster 4	0.174
mistake cluster 5	0.252
mistake cluster 6	0.201
stim surprise cluster 0	0.282
stim surprise cluster 1	0.23
stim surprise cluster 2	0.189
stim surprise cluster 3	0.146
stim surprise cluster 4	0.141
stim surprise cluster 5	0.136
stim surprise cluster 6	0.109
stim surprise cluster 7	0.158
stim surprise cluster 8	0.14
stim surprise cluster 9	0.14
stim congruent cluster 0	0.254
stim congruent cluster 1	0.175
stim congruent cluster 2	0.229
stim congruent cluster 3	0.179
stim congruent cluster 4	0.164
stim congruent cluster 5	0.139
stim congruent cluster 6	0.165
stim congruent cluster 7	0.165
stim congruent cluster 8	0.124
stim congruent cluster 9	0.149
motor init cluster 0	0.307
motor init cluster 1	0.192
motor init cluster 2	0.204
motor init cluster 3	0.289
movement cluster 0	0.234
fback incorrect cluster 0	0.198
fback incorrect cluster 1	0.287
fback incorrect cluster 2	0.214
fback incorrect cluster 3	0.231
fback incorrect cluster 4	0.199

Table 3. Stability scores for shuffled control clusters. Shuffled cluters generated by shuffling region labels of cells before calculating regional similarity scores and identifying groupings with Louvain algorithm. None of these shuffled clusters show stability scores > 0.48 that pass the stability criteria.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

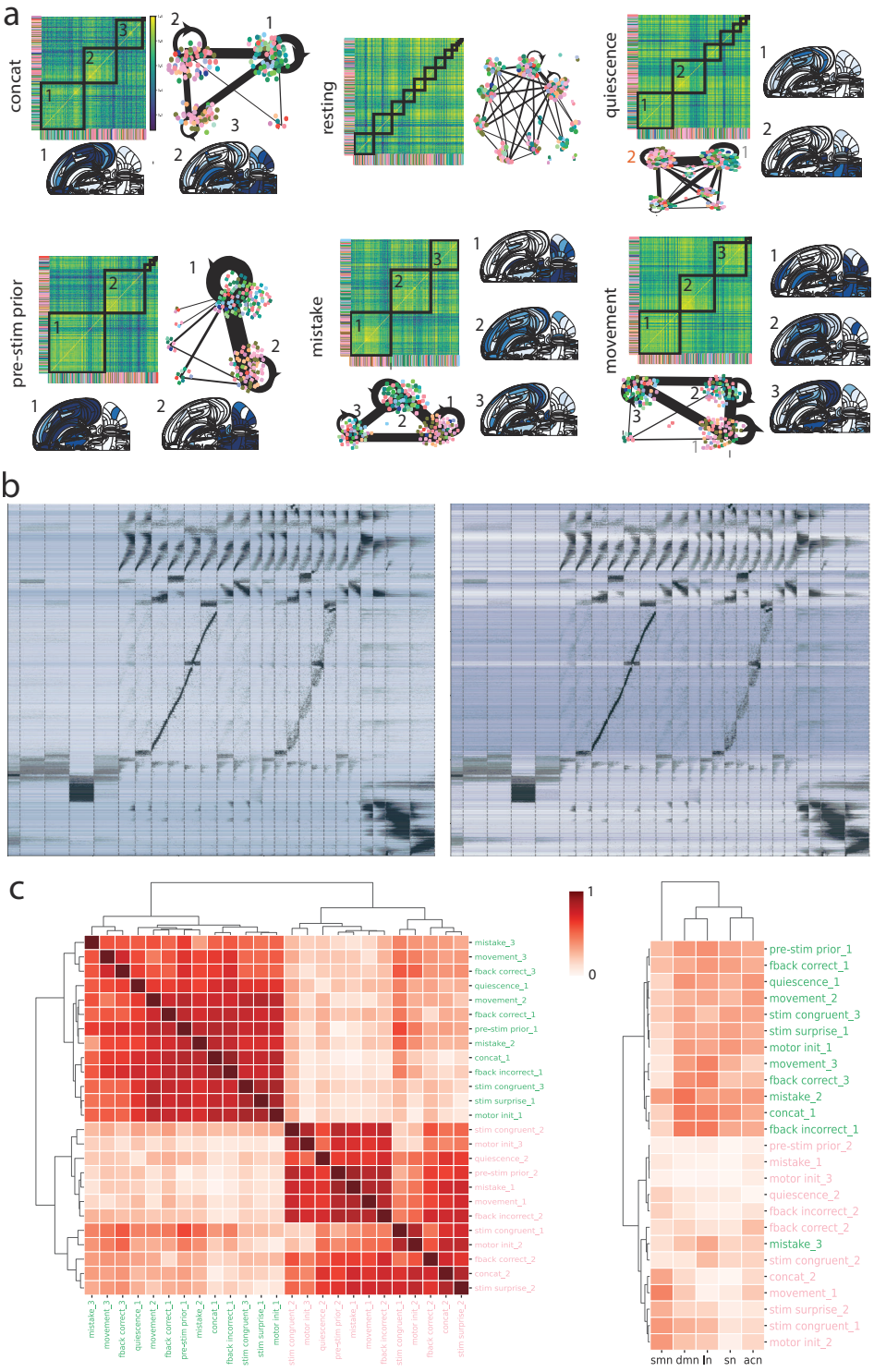


Figure S10. Regional clusters in concat and subset feature vectors a. For completeness, we show the results of Louvain community detection on regional similarity for task phases that have some or all networks which do not pass the statistical significance tests and are excluded from the main figure 6. b. We colored the raster plot (after sorting from Rastermap algorithm) of full concat PETHs with the two clusters identified by Louvain algorithm, using the same color scheme of the Swanson maps (darker blue indicate stronger/ more reliable membership in the cluster). We see a similar pattern of the two networks as in fig.3c and d, reflecting a cortical-subcortical separation. c. For completeness, we show the similarity scores between all stable networks from all task phases, and between these networks and existing fMRI networks. The similarity matrices shown in the main figure 6 are a subset of these networks, and we see the same cortical-subcortical separation pattern.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

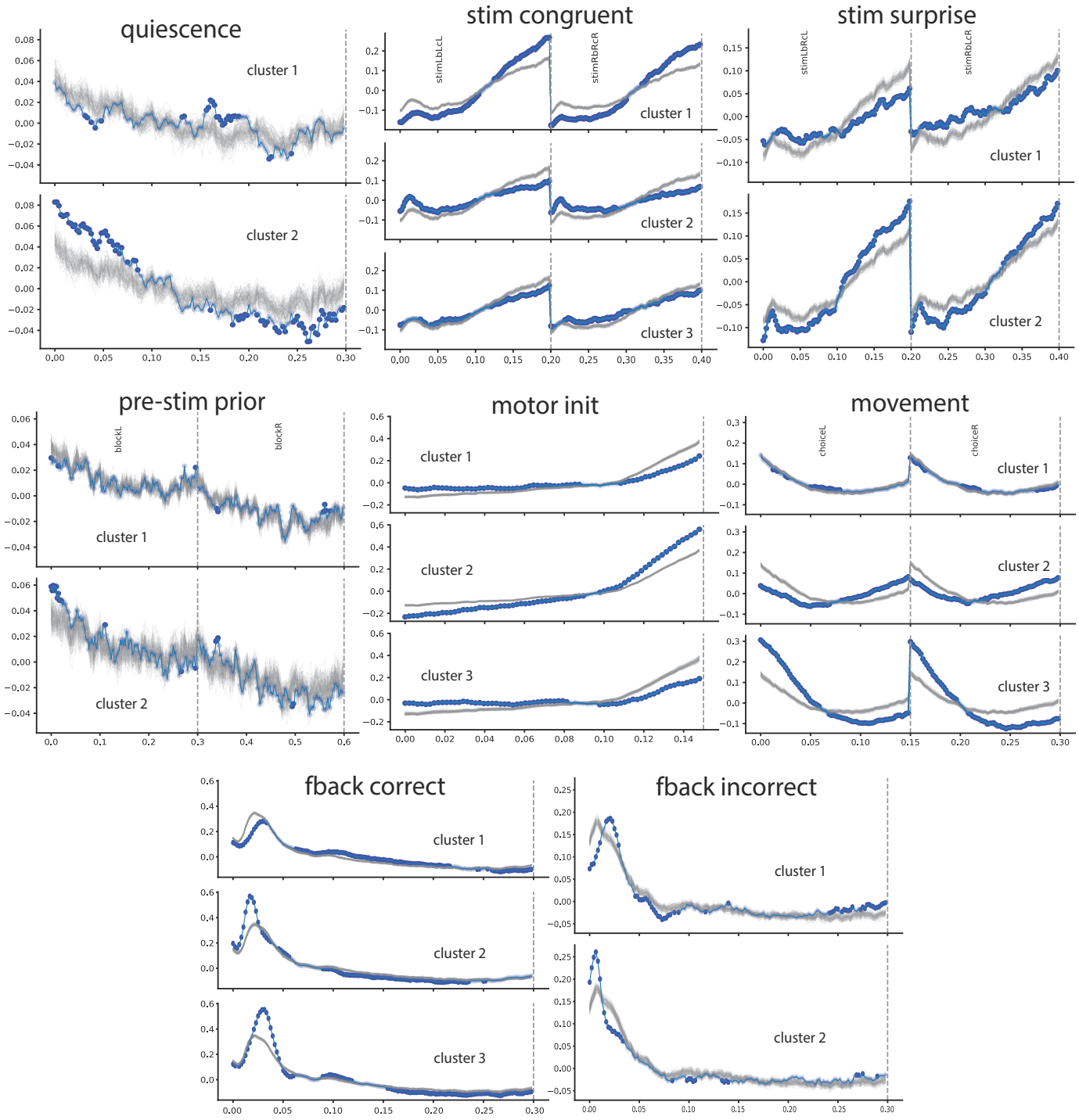


Figure S11. Average PETH of clusters during different task phases True clusters average PETH (blue) compared with controls generated by randomly shuffling the regional membership of each cluster (gray lines). Dark dots on the true data indicates significant difference from the controls.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

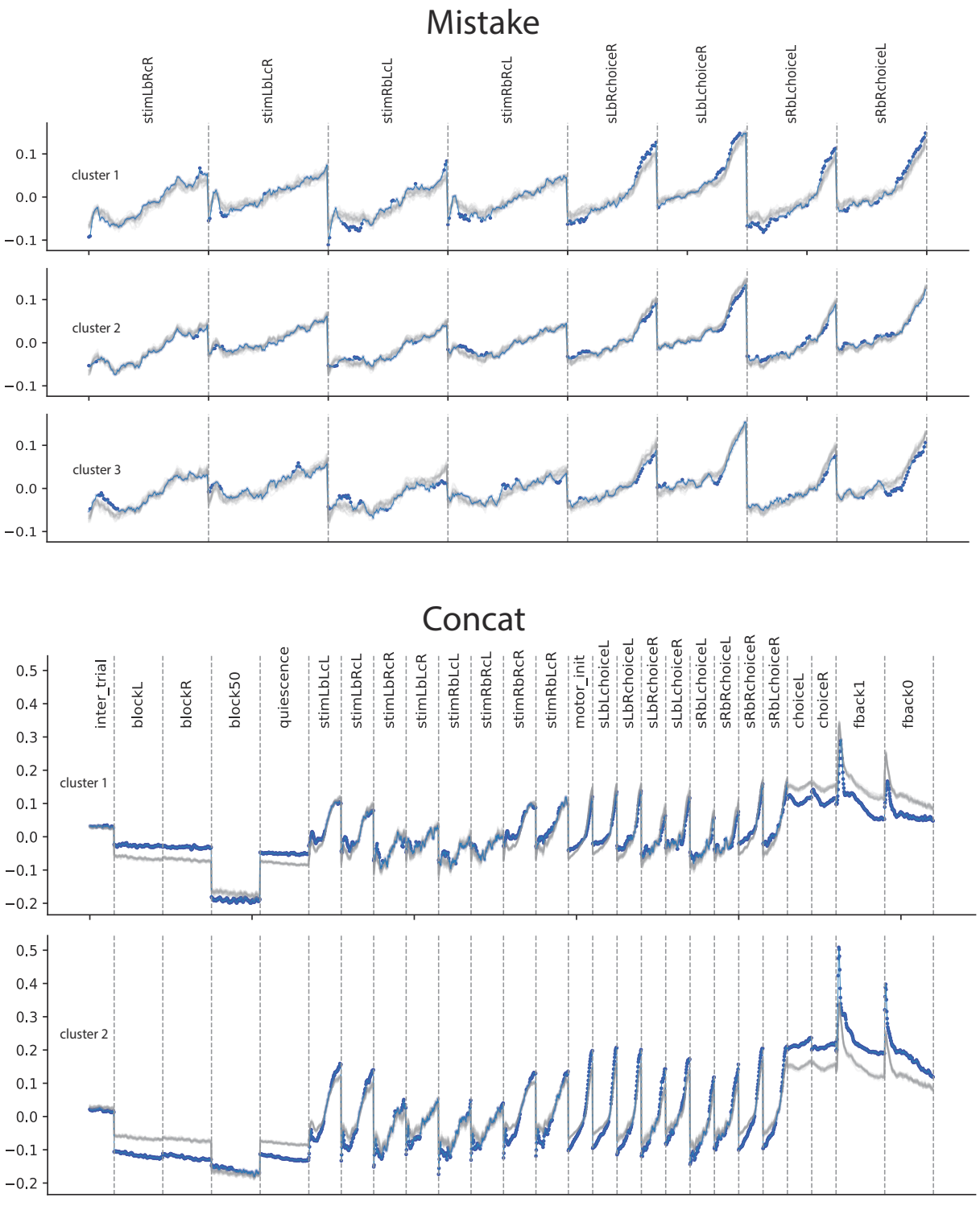


Figure S12. Average PETH of clusters during different task phases (contd.) Same shuffled control analysis as in the previous figure, for the mistake trials and the full concat feature vectors.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

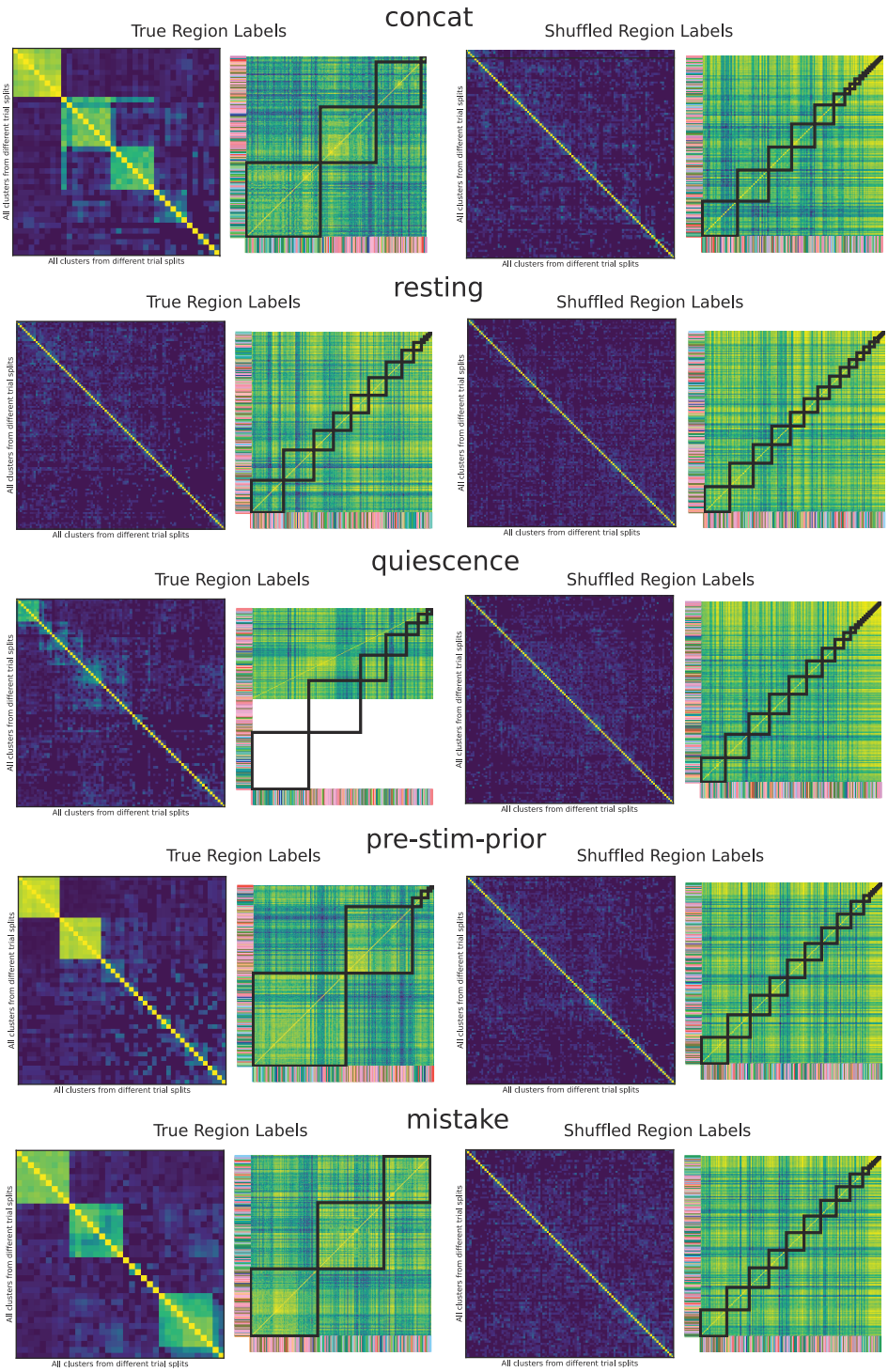


Figure S13. Shuffled controls for Louvain community detection Created by randomly shuffling region labels of all cells before computing the functional similarity scores between regions and running Louvain algorithm to identify clusters. We examine the resulted clusters using these shuffled region labels (right half of each panel) compared to using the true region labels (left half of each panel), by computing cosine similarity between all the clusters generated from randomly selecting a number of trials from each trial type when creating feature vectors. Using the true region labels, Louvain algorithm detects some reliable clusters that are consistent across different shuffles of trials (for most trial phases except resting), whereas controls using the shuffled region labels always result in inconsistent clusters. Similarity scores among clusters using different subsets of trials are shown as big matrices ordered by hierarchical clustering along with dendrograms, and the clusters identified by Louvain algorithm on the feature vectors averaged over all trials are shown in the accompanying smaller matrices.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

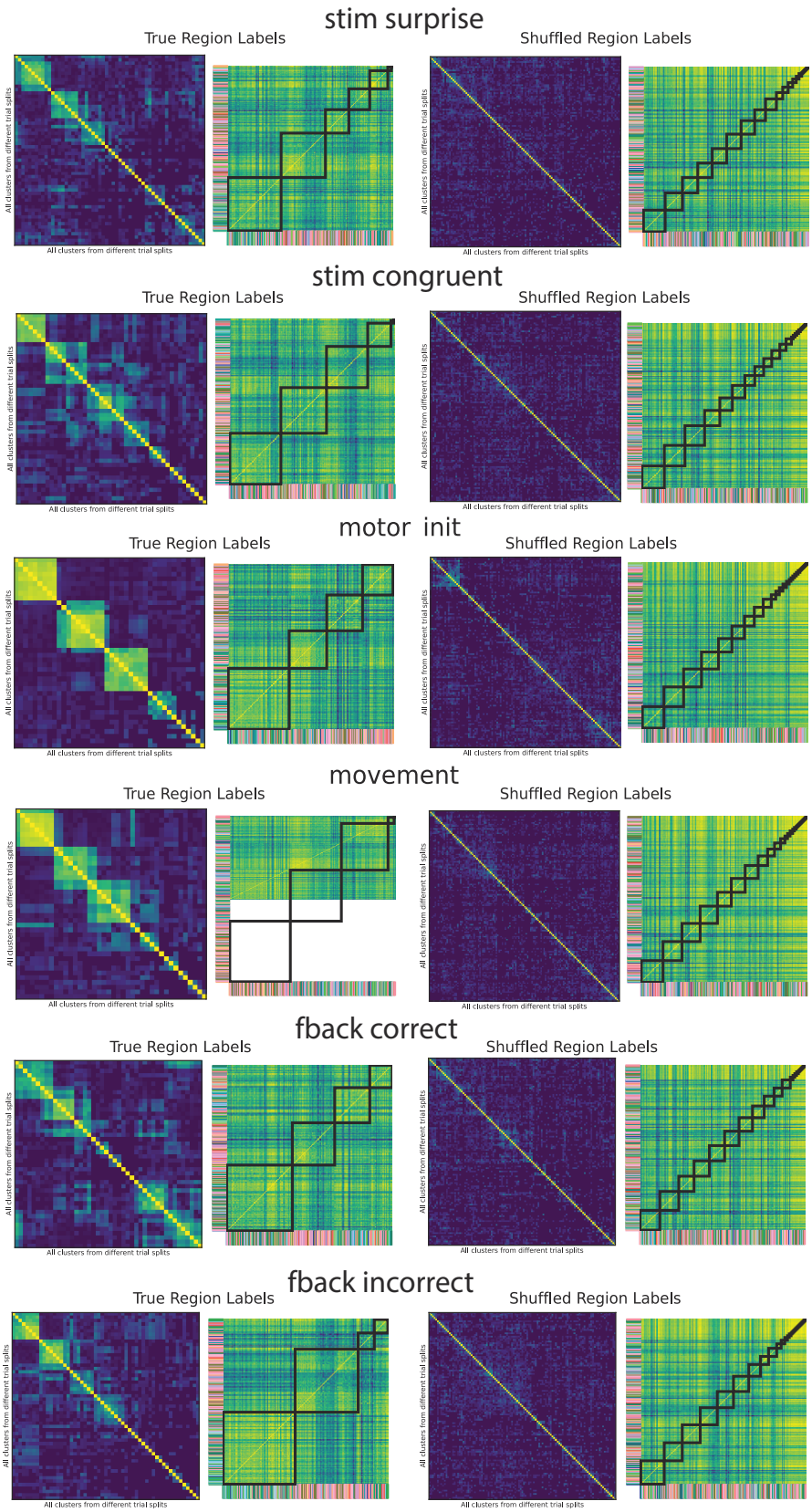


Figure S14. Shuffled controls for Louvain community detection (contd.) For all the remaining task phases/ trial types, the Louvain algorithm detects some reliable clusters that are consistent across different shuffles of trials, whereas controls using the shuffled region labels always result in inconsistent clusters.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



Figure S15. Anatomical hierarchy for all used regions For each of the used regions, at "Beryl" hierarchical level, the parent structures are shown with full names and acronyms, up to "Cosmos" hierarchical level. Find the full anatomical hierarchy tree at the Allen adult mouse brain atlas.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

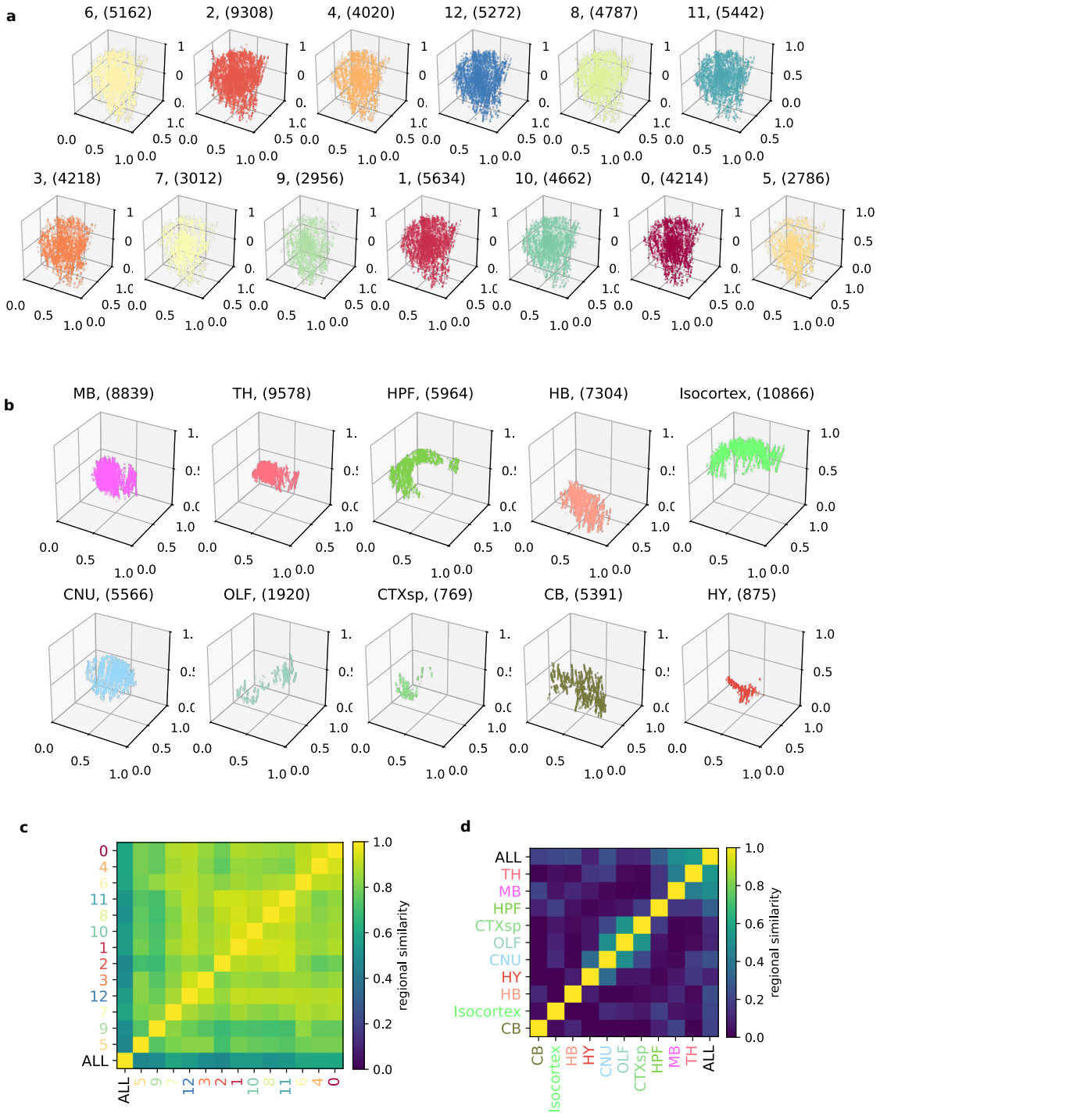


Figure S16. Anatomical xyz position distances for cell partitions controlled by global density a Partition of cells into functional kmeans groups; cluster number in title (number of neurons), note homogeneity in cell numbers and spatial distribution. **b** Partition of cells into anatomical Cosmos regions. **c** Similarity across functional neuron groups (kmeans clusters) after normalizing each smoothed 3D voxel image by the image of all cells (ALL). Note quantitave reduction of kmeans cluster similarity in comparision to Fig.2f, while qualitatively all clusters remain similarly distributed in space. **d** Similarity across anatomical neuron groups (Cosmos) after normalizing each smoothed 3D voxel image by the image of all cells (ALL). Barely any difference in comparision to Fig.2g.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

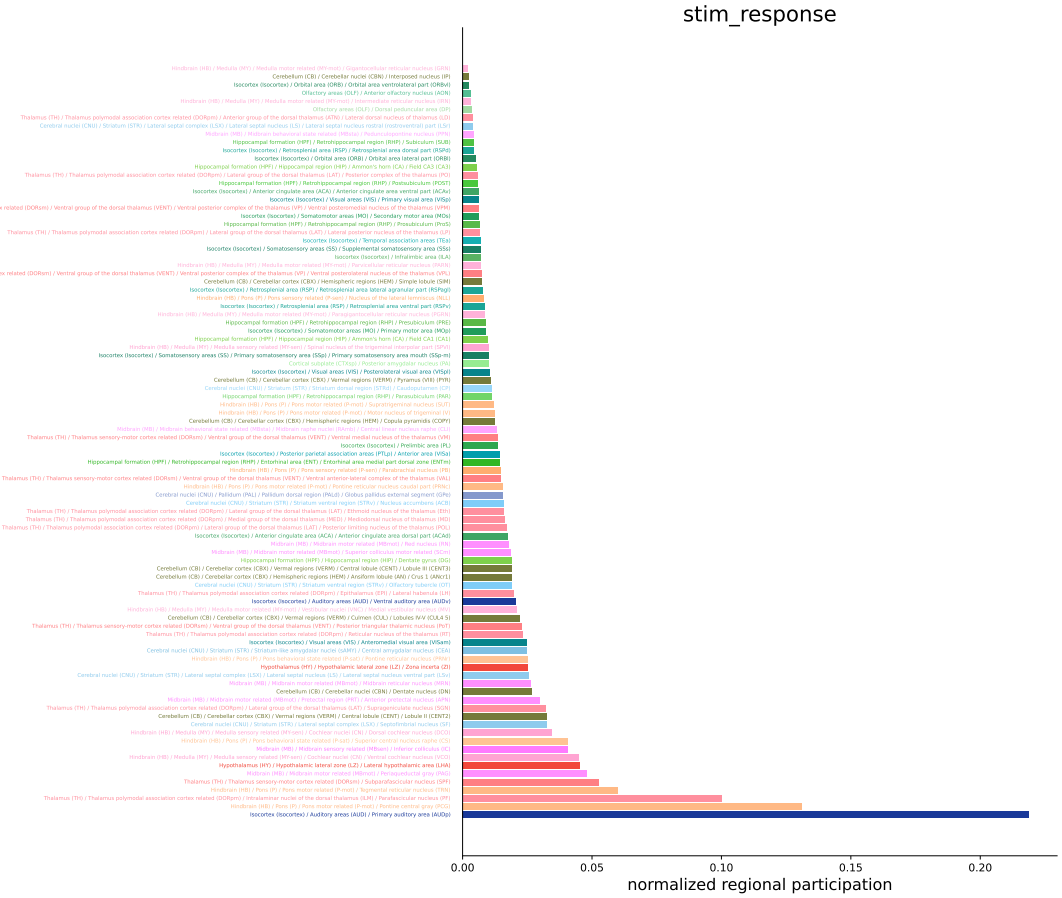


Figure S17. Regional participation histogram in stimulus response cells with the full anatomical hierarchy listed for each region.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

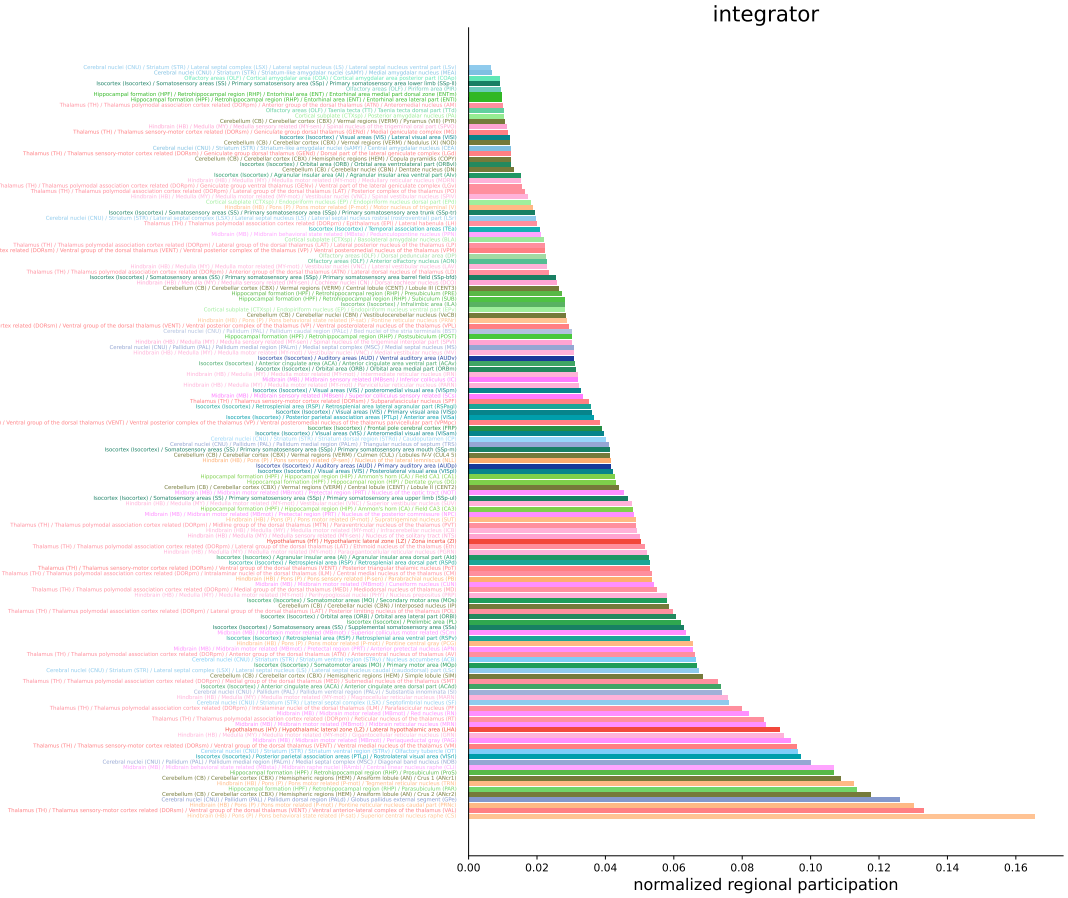


Figure S18. Regional participation histogram in integrator cells with the full anatomical hierarchy listed for each region.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

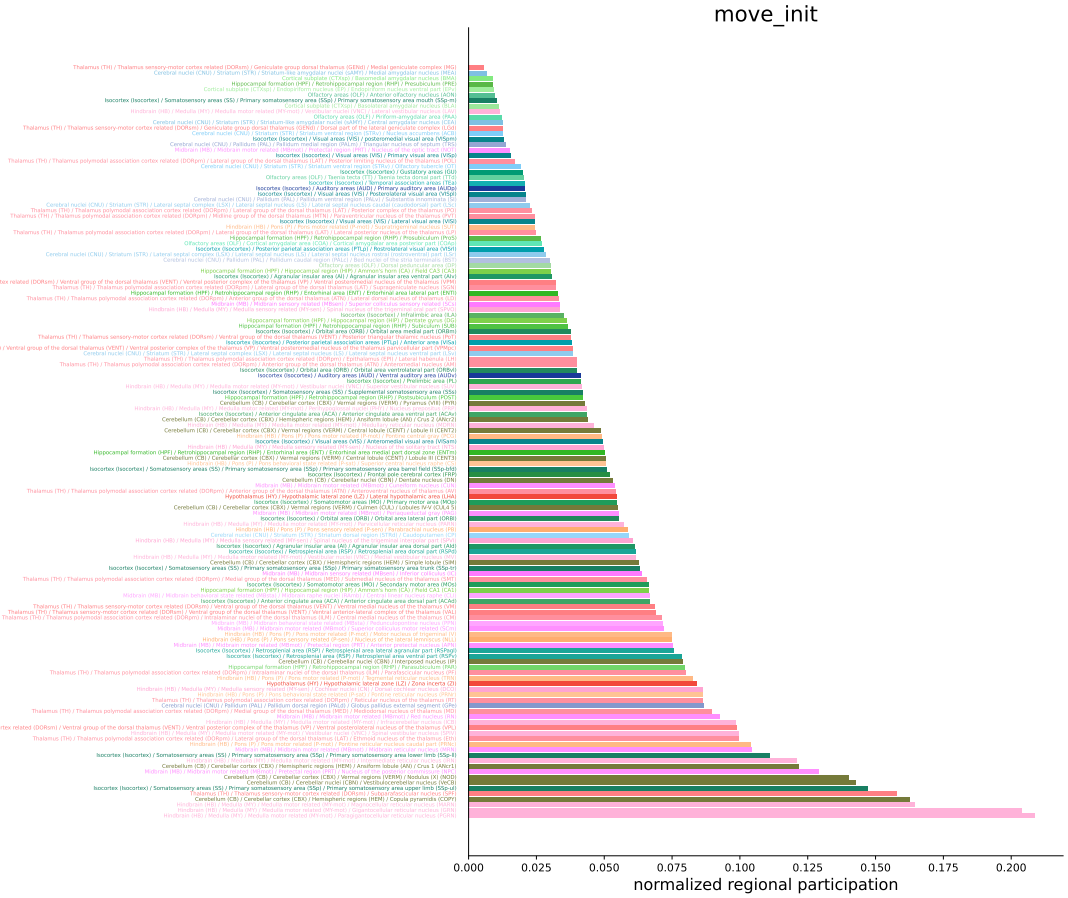


Figure S19. Regional participation histogram in movement initiation cells with the full anatomical hierarchy listed for each region.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

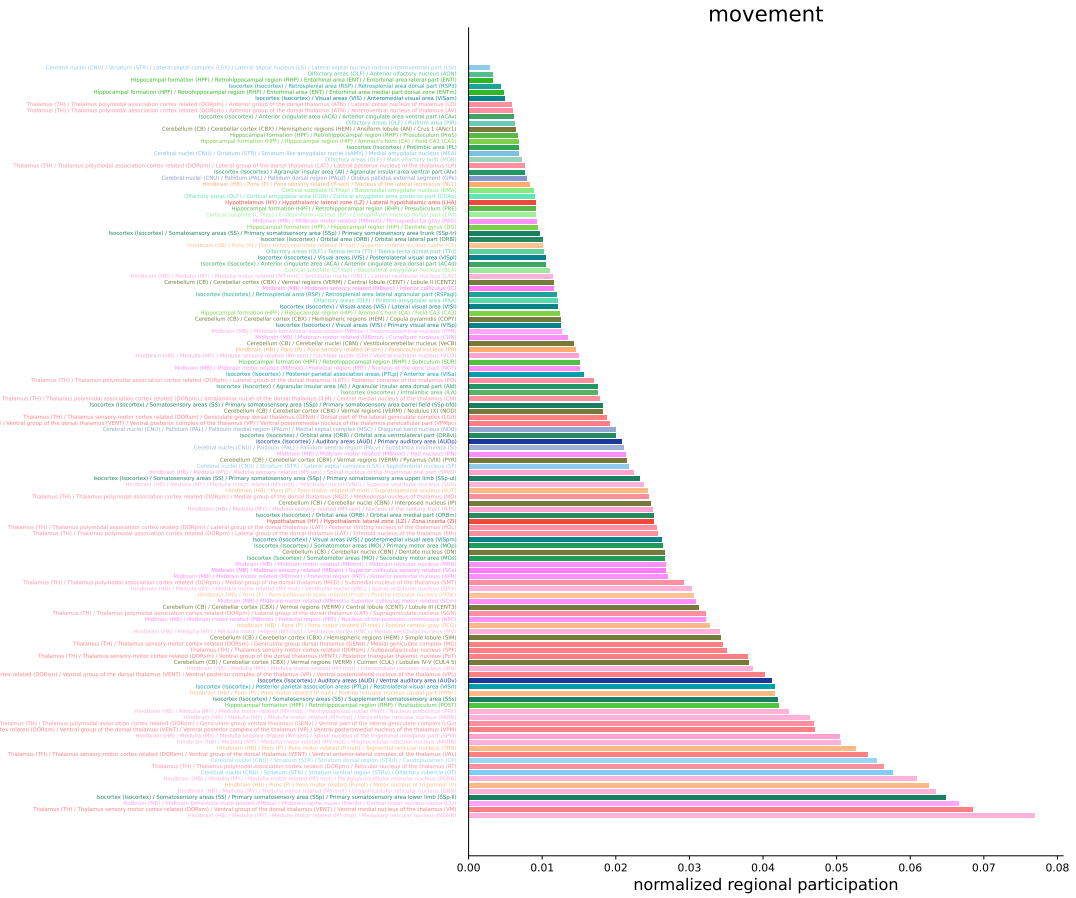


Figure S20. Regional participation histogram in movement cells with the full anatomical hierarchy listed for each region.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



Figure S21. Regional participation histogram in error encoding cells with the full anatomical hierarchy listed for each region.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.30.667641; this version posted August 1, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



Figure S22. Regional participation histogram in sequence cells with the full anatomical hierarchy listed for each region.