Reviewed Preprint
v1 • December 16, 2025
Not revised

Reviewed Preprint
v2 • June 12, 2026
Revised by authors

Neural activity profiles reveal overlapping, intermingled subpopulations spanning area borders in mouse sensorimotor cortex

Sohrab Salimian[1,*], **Harrison A Grier**[1,*], **Matthew T Kaufman**[2,3,4]

[1]Committee on Computational Neuroscience, The University of Chicago, Chicago, United States • [2]Dept. of Organismal Biology and Anatomy, The University of Chicago, Chicago, United States • [3]Neuroscience Institute, The University of Chicago, Chicago, United States • [4]NSF-Simons National Institute for Theory and Mathematics in Biology, Chicago, United States

For correspondence: mattkaufman@uchicago.edu

[*] Equal contributions

Competing interests: No competing interests declared

Funding: See page 31

Reviewing editor: Brice Bathellier, Centre National de la Recherche Scientifique, France

eLife Assessment

This **fundamental** study provides **compelling** evidence for the functional segregation of the sensorimotor cortex into precisely delineated areas, and highlights a rapid transition in functional properties at the boundaries between these areas. This result further confirms and extends recent work on the diversity of neural response specificities across cortical areas in the context of complex behavioral tasks. This work will be of interest to neuroscientists studying sensory-motor functions.

https://doi.org/10.7554/eLife.109240.2.sa3

Abstract

Cortical control of movement is a distributed computation spanning multiple densely-interconnected regions. Although we have rich anatomical atlases and a coarse understanding of how function maps to areas and subregions, we lack a detailed account of how behaviorally-relevant activity is organized across the cortical sheet. Here, we trained head-fixed mice to perform a 15-target reach-to-grasp task while we performed cellular-resolution, two-photon calcium imaging across five regions of sensorimotor cortex (>39,000 layer 2/3 neurons). We characterized each neuron's trial-averaged peri-event activity with interpretable metrics and mapped these response properties across areas, revealing large-scale spatial structure. Neuronal response profiles often shifted abruptly at anatomical borders: motor areas showed sharper tuning and more linear relationships with target location, whereas somatosensory areas displayed more heterogeneous response patterns. Neural response properties also differed according to somatotopic representation. Nonlinear dimensionality reduction of the neural feature matrix revealed that areas varied in their average response profiles, but that areas did not have well-separated feature distributions; instead, each area contained subpopulations. Neurons in each subpopulation had characteristic response profiles and were distributed across multiple cortical areas. The spatial distributions of the subpopulations overlapped, with neurons from different subpopulations salt-and-pepper intermingled in the overlap zones. Together, these results describe novel activity structure across sensorimotor cortex and identify several distinct but spatially-overlapping subpopulations with characteristic activity patterns during reach-to-grasp behavior.

Introduction

Controlling movement is a critical function of the brain, and a sizable fraction of mammalian neocortex is involved in this task (Currie et al. 2022; Li et al. 2024; Muñoz-Castañeda et al. 2021). Moving in dynamic environments requires processes ranging from trajectory planning

and pattern generation to error monitoring and feedback integration. Because of the importance of movement control, a wide variety of approaches have been used to study motor and somatosensory areas: from examining neural biophysics and connectivity, to determining neurons' direct influence on movement, to relating activity of neurons in relevant subregions to behavior. However, to date the field has not systematically surveyed activity across sensorimotor cortex at single-cell resolution during behavior.

Anatomical mapping has formed the foundation of our understanding of the biological pieces of the brain's puzzle, and has classically leveraged cytoarchitecture and projection tracing to describe how structure varies across the cortical sheet. These approaches support the view that cortex is divided into many areas that serve different functions, while remaining strongly interconnected (Brodmann 2005; Majka et al. 2021; Yeterian et al. 2012) especially in the mouse (Gămănuț et al. 2018; Harris et al. 2019; Oh et al. 2014; Winnubst et al. 2019; Zingg et al. 2014). More recent genetic tools have complemented these methods with finer-grained insight into neuron-type distributions and connectivity, providing additional distinctions between parts of the brain (DeNardo et al. 2015; Muñoz-Castañeda et al. 2021; Wang et al. 2020) and revealing multiple parallel circuits within brain areas (Carmona et al. 2024; Park et al. 2022; Ueno et al. 2018). Subpopulations with distinct projection patterns and layer-specific connectivity suggest the existence of multiple functions coexisting within both somatosensory (Kerr et al. 2007) and motor areas (Carmona et al. 2024; Hausmann et al. 2022; Hira et al. 2013b; Muñoz-Castañeda et al. 2021; Zingg et al. 2014). Moreover, strong projections to subcortical, thalamic, and spinal targets originate from subregions of these cortical areas, some of which closely abide by anatomical borders determined by other methods (Carmona et al. 2024; Jeong et al. 2016) while others cross them (Hausmann et al. 2022; Ueno et al. 2018).

To relate activity to behavior more directly, other approaches have instead correlated neural activity with self-generated behavior. Large-scale recordings with simple behaviors have suggested that motor signals are widespread across cortex (Musall et al. 2019; Steinmetz et al. 2019; Stringer and Pachitariu 2019; Wang et al. 2023b), while studies with slightly more challenging behaviors have shown a greater proportion of motor signals in areas closer to the motor or sensory periphery (Li et al. 2024; Wang et al. 2023b). However, the details of what motor signals reside where have not been delineated. In an even more challenging motor task, neural populations in primary motor cortex (M1) and the primary somatosensory forelimb area yield similarly strong kinematic decoding of many tracked joint angles (Grier et al. 2026). These findings describe a distributed system, where signals are widely shared and where areas' functional distinctions remain unresolved despite their sharp differences in connectivity with the periphery (Carmona et al. 2024; Ueno et al. 2018). A key question then remains: how are the activity patterns of individual cortical cells organized across sensorimotor cortex during behavior, and how does this relate to previous anatomical mapping?

To frame possible outcomes, consider that single neuron responses can vary along many dimensions. Cells could differ according to which movements or time periods they are recruited for (tuning), what movement parameters their activities reflect (encoding), or how their responses are structured across different movements (e.g., nonlinear encoding structure). Further, differences in these response properties across cells could be distributed over the cortical sheet in a variety of ways. Cells could form distinct "categories" or clusters that are spatially well-aligned to the boundaries of anatomically defined regions. Or, categories of neurons might span area boundaries in spatial footprints that do not relate obviously to area boundaries, and that either abut or overlap. At a fine-grained scale, cells with similar responses could be physically located near one another as in primate and feline visual cortex, or similarly-responsive neurons might be salt-and-pepper intermingled as seen in rodent visual cortex or in primate motor cortices during reaching behaviors.

Here, at single-cell resolution, we densely mapped single neuron activity across mouse sensorimotor cortex during a reach-to-grasp behavior to systematically capture the spatial distribution of tuning properties. We used two-photon calcium imaging to record from over 39,000 neurons as mice performed a 15-target reach-to-grasp task. We focused on five areas involved in

this behavior: secondary motor cortex (M2), M1, and the primary sensory areas corresponding to the forelimb (S1-fl), hindlimb (S1-hl), and trunk (S1-tr). Our analyses emphasized the format of how neurons encoded information, rather than correlating activity with the behavior directly. Some tuning properties related to response onsets and peak timing closely followed anatomical borders. Other properties were more closely related to somatotopy, such as tuning sharpness, response duration, or the linearity of responses with respect to target location. Using bottom-up approaches, we were also able to identify subpopulations that shared response features, which were salt-and-pepper intermingled within anatomically-defined areas but had characteristic spatial footprints across cortex that spanned multiple areas. Together, these results clarify the organization of the response properties of neural populations at multiple scales, including identifying previously-unknown, overlapping and intermingled area-spanning subpopulations.

Results

Dense sampling of sensorimotor cortex reveals heterogeneous tuning

We implanted cranial windows covering sensorimotor cortex (Figure 1A-B) in six male mice expressing GCaMP8s in most cortical pyramidal cells (Zhang et al. 2023). Dense sampling of layer 2/3 neurons was achieved using two-photon calcium imaging across 98 fields of view (FOVs) spanning the forelimb, hindlimb, and trunk representations (Figure 1D). FOVs were aligned to the Allen Mouse Brain Common Coordinate Framework version 3 (Allen CCF) (Wang et al. 2020) using a combination of anatomical landmarks and widefield imaging with paw vibration mapping (Methods; Figure 1B). A total of 39,398 neurons were extracted across the five relevant regions in the sensorimotor cortex (Figure 1E). A larger number of neurons were imaged in M2 and M1 (31,453 neurons) than in somatosensory areas (7,945; Figure 1C).

Behavior

While neural activity was acquired the mice performed a trained, motorically-challenging, delayed reach-to-grasp-to-drink task (Methods). The task was a modified version of the Galiñanes and Huber (2023) multi-target reaching task. Mice were head-fixed and held two paw rests to initiate a trial (Figure 2A,B). A trial began with a spout moving into one of 15 positions, selected randomly on each trial from a concave grid arranged around the animal's snout. After a variable delay period a non-target-specific auditory Go cue ("cue") was played and a water droplet released, and the mouse could use its right paw to reach toward the target and retrieve the water reward. Stereo high speed camera data enabled markerless tracking (DeepLabCut; Mathis et al. (2018)) offline. Animals were successful at contacting the target on 97.1% of trials, and an average of 373 successful trials were collected per session (264-450; 39,229 successful trials total in the dataset).

As a simple characterization of the behavioral kinematics, single-trial paw centroid trajectories were plotted and colored by their corresponding target position (Figure. 2C and Figure 2-figure supplement 1A). Two relevant features were readily apparent. First, the paw centroid location was strongly target-specific near the spout, demonstrating that reaches were in fact targeted. Second, a looping trajectory near the spout was typically present, which corresponds with repeated reaches to collect all the water from the spout. This was due to a combination of our custom-shaped spouts, which were designed to cause the water to cling, and our use of touch sensors which encouraged strong contact with the spout itself (Methods). Although variability was present in both reaction time (Figure 2D) and movement time (Figure 2E), as is typical in reaching data, both intervals were generally short and indicated that animals were highly engaged in the task. To assess the extent to which mice aimed their reaches to each target, we computed the mean Euclidean distance across trials within target and between targets (Figure 2-figure supplement 1B). Reaches made to the same target were more similar to one another than were reaches to different targets, suggesting that the animals intentionally reached to each spout location rather than generically reaching out and searching for contact. We do not have data on how the mice localized the spout on each trial, but it could potentially have been by sniffing

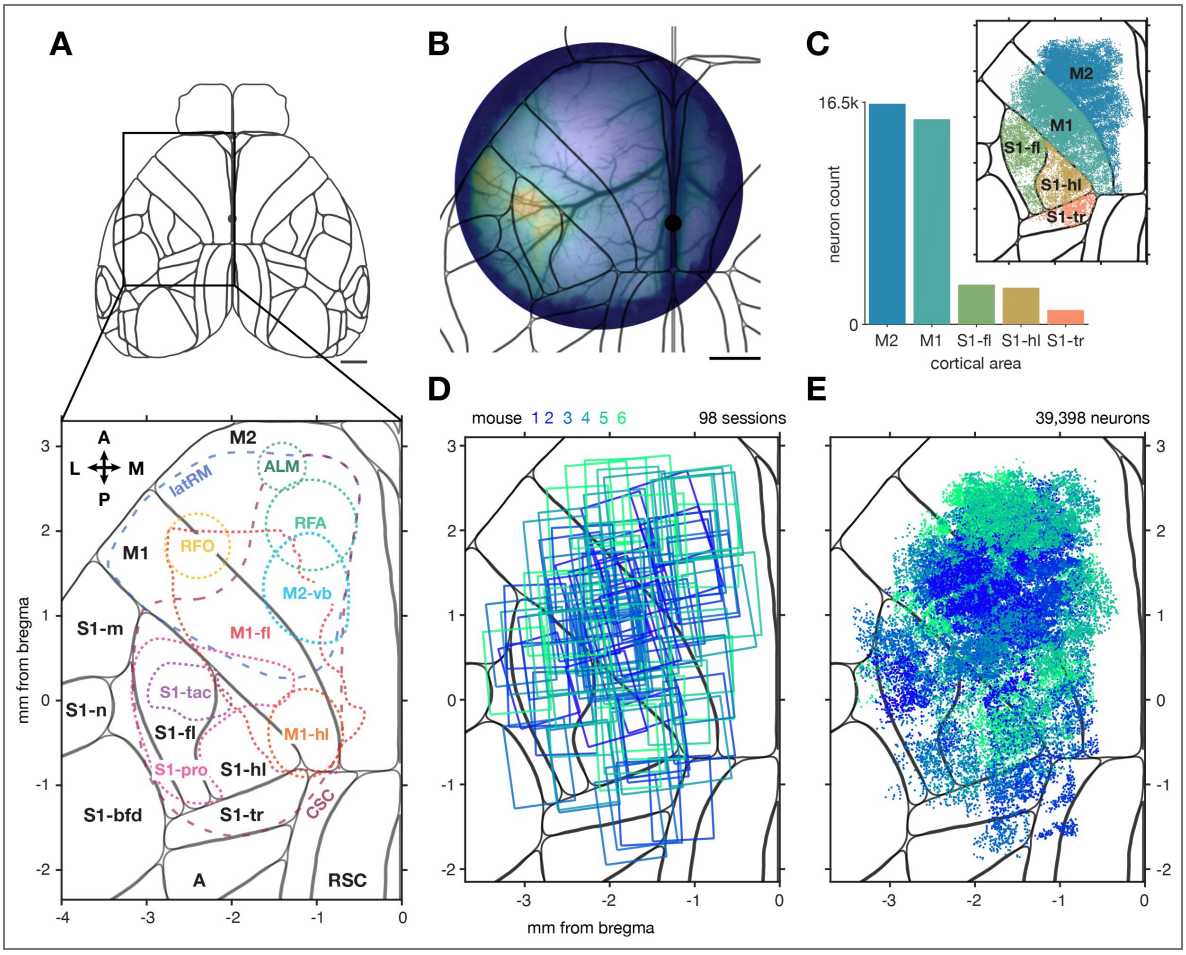


Figure 1. A dataset of densely sampled activity across mouse sensorimotor cortex.

(A) *Top*, dorsal view of the Allen CCF, with the area investigated here indicated by a box. *Bottom*, summary of several previously-identified subareas of mouse sensorimotor cortex (see Methods for descriptions, alignment, and sources). Orientation legend: A, anterior, P, posterior, L, lateral, M, medial. (B) Widefield calcium imaging in a 5 mm cranial window superimposed on the Allen CCF to show location. Activation elicited by paw vibration (Methods). (C) Counts of neurons imaged for each area, with identities colored in inset. (D) Field of view locations of the 98 sessions of two-photon imaging collected in layer 2/3 for 6 mice (colors). (E) Locations of the 39,398 neurons extracted from all sessions. Mouse identity colored as in D.

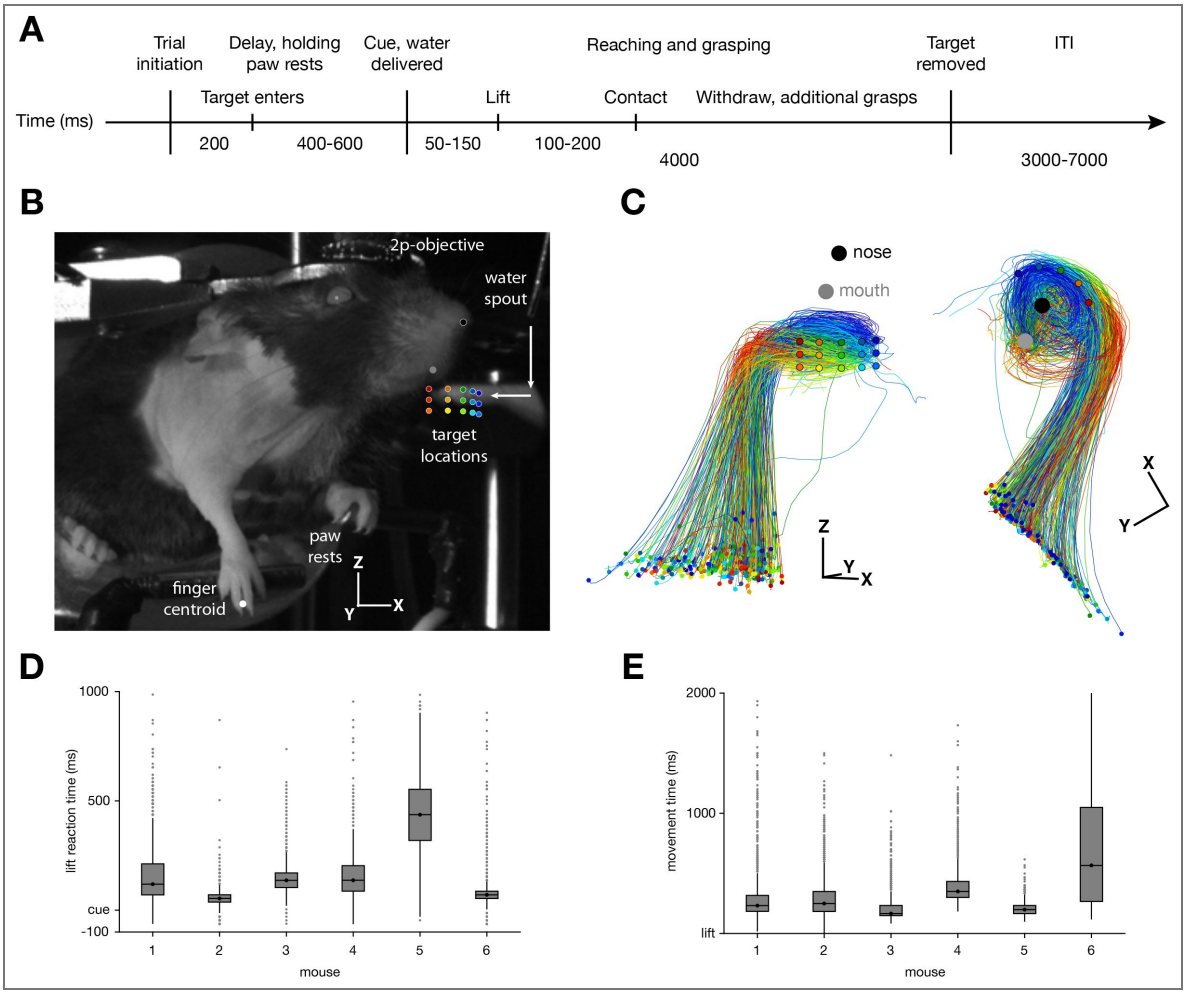


Figure 2. Mice performed delayed reach-to-grasp-to-drink movements to 15 distinct targets.

(A) Task timeline. ITI, inter-trial interval. (B) Infrared image of the mouse during the inter-trial interval from one of two high-speed cameras. The 15 possible target locations illustrated with colored dots. (C) Finger centroid trajectories locked to lift onset from a frontal (left panel) and dorsal (right panel) view, illustrating kinematic separation by target. Target locations and average nose and mouth locations superimposed over centroid traces for clearer visualization. Coordinate axes show 3 mm scale bars and correspond to the axes in B. (D) Box plots of lift reaction times (RTs) for each mouse. Boxes show interquartile range (IQR) with a line for median; whiskers show 1.5 times the IQR; dots show outliers. Negative RTs indicate reaches before the Go cue, and were excluded from further analysis. (E) Time between lift and spout contact across mice. Three consecutive sessions per animal shown for both D and E.

(Galiñanes et al. 2018), whisking, or detecting vibrations from the motors that positioned the target. A fuller characterization of how neural activity relates to the kinematic output will be detailed in forthcoming work.

Neural data visualization

To begin investigating the neural activity, we identified which neurons modulated their firing rates during the task. To do so, we computed a modulation score for all cells using a variant of the ZETA method (Methods; Montijn et al. (2021)), which asks whether the activity for a neuron changes over time consistently in relation to any given target. Modulated neurons were common and scattered across all five regions of interest in the sensorimotor cortex (Figure 3A). To quantify their spatial distribution, we computed the fraction of neurons that were modulated across the cortical sheet. The five areas of interest varied in the fraction of neurons that were modulated: M2 had 14%, M1 had 23%, S1-fl had 30%, S1-hl had 25%, and S1-tr had 27% ($p < 10^{-16}$, Chi squared test for homogeneity, cue-aligned; between-area pairwise two proportion z-test in Figure 3-figure supplement 1B; Methods). Given that these areas have known spatial organization within them and structure was apparent by eye in the spatial scatterplot of modulated neurons (for example, the less modulated band along the M1/M2 border in Figure 3A), we then produced a high-resolution map of the fraction of neurons that were modulated by location (Figure 3B). Neurons in two zones were more often modulated than elsewhere: a wide strip along the border between S1 and M1, and an anteromedial patch of M2. The probability of a neuron being task-modulated was highest around the border between M1 and S1-fl, corresponding to the center of the forelimb portion of M1 and the proprioceptive region of S1-fl (Alonso et al. 2023). Probability of modulation was particularly low in a wide strip along the border between M1 and M2.

Modulation data are shown here computed with neural activity aligned to the Go cue. When aligning data to the paw lift (“Lift”) or spout contact (“Contact”), a mostly-overlapping set of cells were identified as modulated (Figure 3C, Figure 3-figure supplement 1). Although more cells were identified as modulated when using the Lift or Contact event alignments, all results in this work were similar across all three possible alignments. Results are shown through the rest of this work aligned to Cue because that alignment best preserved some unique activity structure in M2 (noted below) while preserving all other findings.

We next visualized the tuning profiles of individual neurons through peri-event time histograms (PETHs). Five representative example neurons for each of the five areas of interest are shown (Figure 3D-H). One striking feature of the neural responses was that neurons were often strongly tuned, with large differences in the activity associated with different targets, despite the fact that the targets were relatively close together in physical space. Several systematic differences between the neurons in each area were also prominent, both in these examples and in our qualitative assessment of the recorded population overall. Neurons in both M2 (Figure 3D) and M1 (Figure 3E) were often strongly tuned, with firing rates ranging from no measurable activity for some targets to strong and reliable responses to other targets. M1 neurons typically exhibited briefer responses than M2 neurons, while M2 neurons were often more sharply tuned – that is, strongly active for a smaller number of targets. S1-fl (Figure 3F) typically exhibited the briefest responses, with a consistent peak of activity for most or all targets. S1-hl (Figure 3G) and S1-tr (Figure 3H) neurons had a large variety of response profiles, frequently showing sustained and temporally complex responses and having activity peaks at substantially different times for reaches to different targets. While examples of almost any response profile could be found in each of the recorded areas, these observations reflect systematic differences in the distributions of responses across these areas, quantified below.

Onset of neural activity varied with somatotopy and subregion

To quantify the overall temporal profile of single neurons, we measured when activity for each neuron ramped up. To compute these onset times, we first averaged the activity for each modulated neuron across all trials, creating a single grand-mean trace per neuron, then found the time to half-max. We next asked whether onset times differed by anatomical area (Figure 4A).

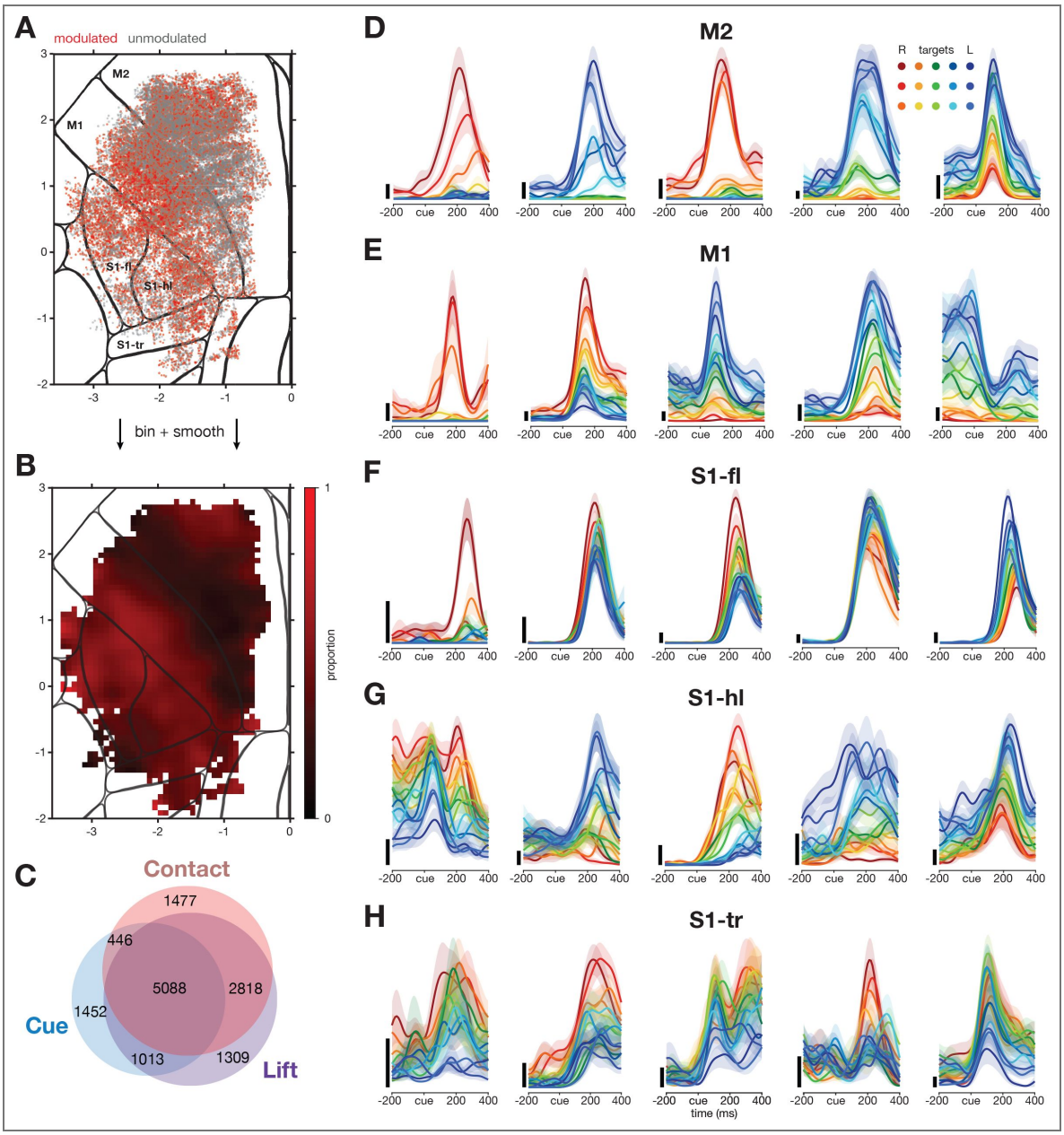


Figure 3. Neurons in mouse sensorimotor cortex exhibited heterogeneous tuning profiles.

(A) Extracted neuron locations superimposed on the Allen CCF. Red indicates a neuron that was modulated for any of the three locking events according to our statistical test (Methods), gray indicates a neuron that was not. (B) Binned and Gaussian-smoothed (s.d. 150 μm) map derived from A. (C) Venn diagram depicting the number of modulated cells for each event alignment. (D-H) Peri-event time histograms (PETHs) of five example neurons in each area. Successful trials to each target were averaged and smoothed. Colors as in Figure 2B. Shaded regions, SEMs; scale bars, 10 events/second. Neurons were chosen to be clear and representative examples of response profiles observed in each area.

These distributions over neurons revealed clear differences in the overall profile of activation: early onsets were more prevalent in S1 trunk and hindlimb regions, perhaps due to activity related to the animal stabilizing itself even if the neurons became more active later; then M2, and finally S1-fl and M1. Nevertheless, each area contained neurons activated at any given time in the trial.

Next, we leveraged our dense spatial sampling to build a map of the time of activation across cortex. The temporal onset metric was computed for each neuron, neurons were binned into pixels, then this map was lightly smoothed (Methods). The resulting temporal onset map revealed a number of sub-area features (Figure 4B). The earliest activations, occurring during the delay (after the target had moved in but well before the Go cue), were in S1 at the juncture of the trunk, hindlimb, and forelimb regions; and in the portions of M1 and M2 that correspond somatotopically to the hindlimb. As suggested by the area-level result above, these zones presumably reflect the animal preparing its body to support the reach. Shortly after, there was activation in a small zone of S1 that corresponded closely with the tactile subregion of the forelimb region, and in the vibrissal subregion of M2 (see Figure 1A; Esmaeili et al. (2021); Mayrhofer et al. (2019)). This latter activation might relate to the animal whisking, possibly to find the target, or because the incoming spout simply triggers a whisking response, or because the animal is engaged (Steinmetz et al. 2019) and anticipating movement. Next, activation spread through much of M2, part of the forelimb portion of M1, and into other parts of S1; and then finally to more lateral parts of M1 and M2 that relate to the face and oromanual feeding (Figure 1A; (An et al. 2025)).

Activity profiles of single neurons varied systematically across sensorimotor cortex

To quantify the heterogeneity of tuning profiles across single neurons, we developed several summary metrics that captured the most salient variations in observed properties across our inspection of thousands of PETHs. This provides greater interpretability than using unsupervised methods such as Principal Component Analysis (PCA), and is consistent with the practice of feature engineering to extract known-meaningful structure from data where more generic methods fail to capture this structure (e.g., in spikesorting; Caro-Martin et al. (2018), Lee et al. (2021), Lewicki (1994)). Brief intuition is given for each metric below, with complete explanations of each in the Methods.

Response duration

To capture whether a neuron was active only briefly or for a more extended time, we measured the width of the autocorrelation for each neuron. Autocorrelations were averaged across all targets for which the neuron was modulated. Larger values indicate that a neuron remained active for longer. This metric was highest in the posterior regions S1-tr, S1-hl, and the portions of M1 and M2 corresponding to the hindlimb (Figure 5A).

Peak time variation

To quantify whether a neuron's firing peaked at the same time for every target or varied by target, we found the peak firing rate of the response to each target, then computed the standard deviation of these peak times across targets. This value is therefore higher if the peak time varied and nearly zero if the timing was consistent. Notably, this measure correlated substantially with overall signal-to-noise ratio of a neuron's PETH (Spearman's ρ=-0.53; Methods), and thus partly measures trial-to-trial variability, not just true peak timing variability. This metric was quite low in M1, indicating highly consistent timing of the activity peak (and reliable responses), and was highest in the posteromedial part of M2 (presumably corresponding to the hindlimb representation) and the posterior tip of S1-hl (Figure 5B).

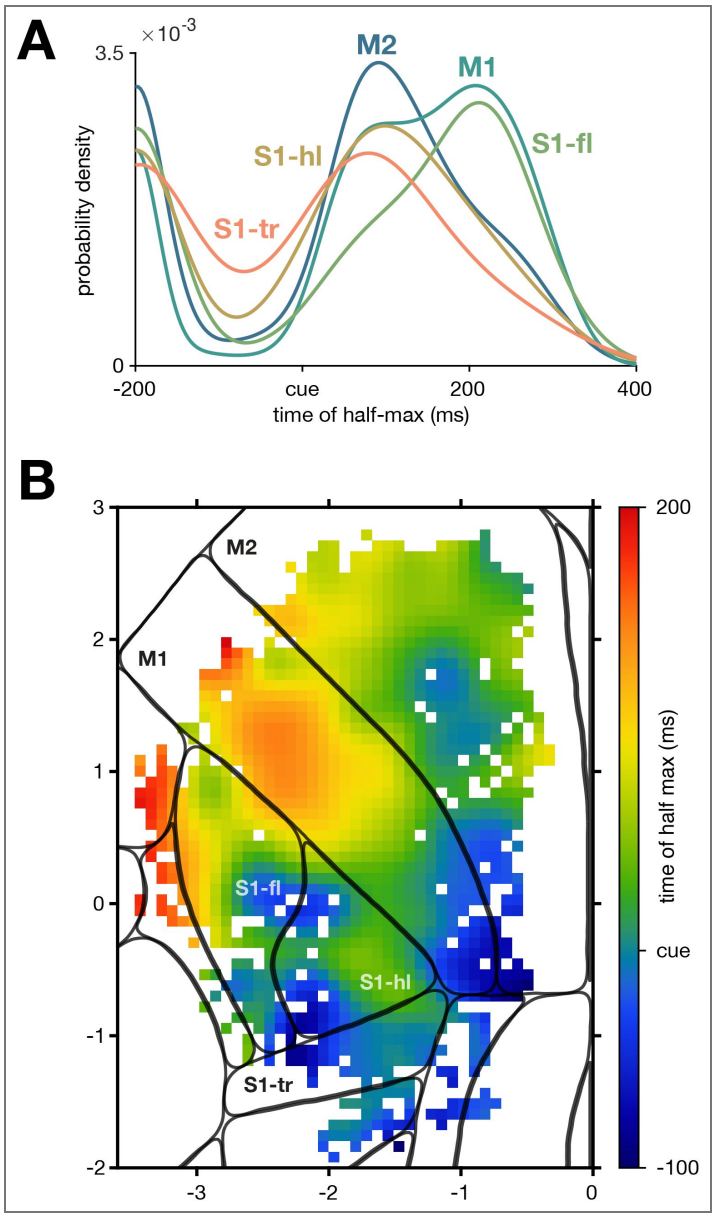


Figure 4. Onset of neural activity across sensorimotor cortex followed both area borders and a somatotopic organization.

(A) Smoothed histogram of time to half-max activation for neurons by anatomical area. (B) Time of half-max activation map. The metric was computed on individual neurons, neurons were binned into pixels, then the map was Gaussian smoothed (s.d. 150 μm). Colorbar ticks indicate range of values in binned and smoothed map.

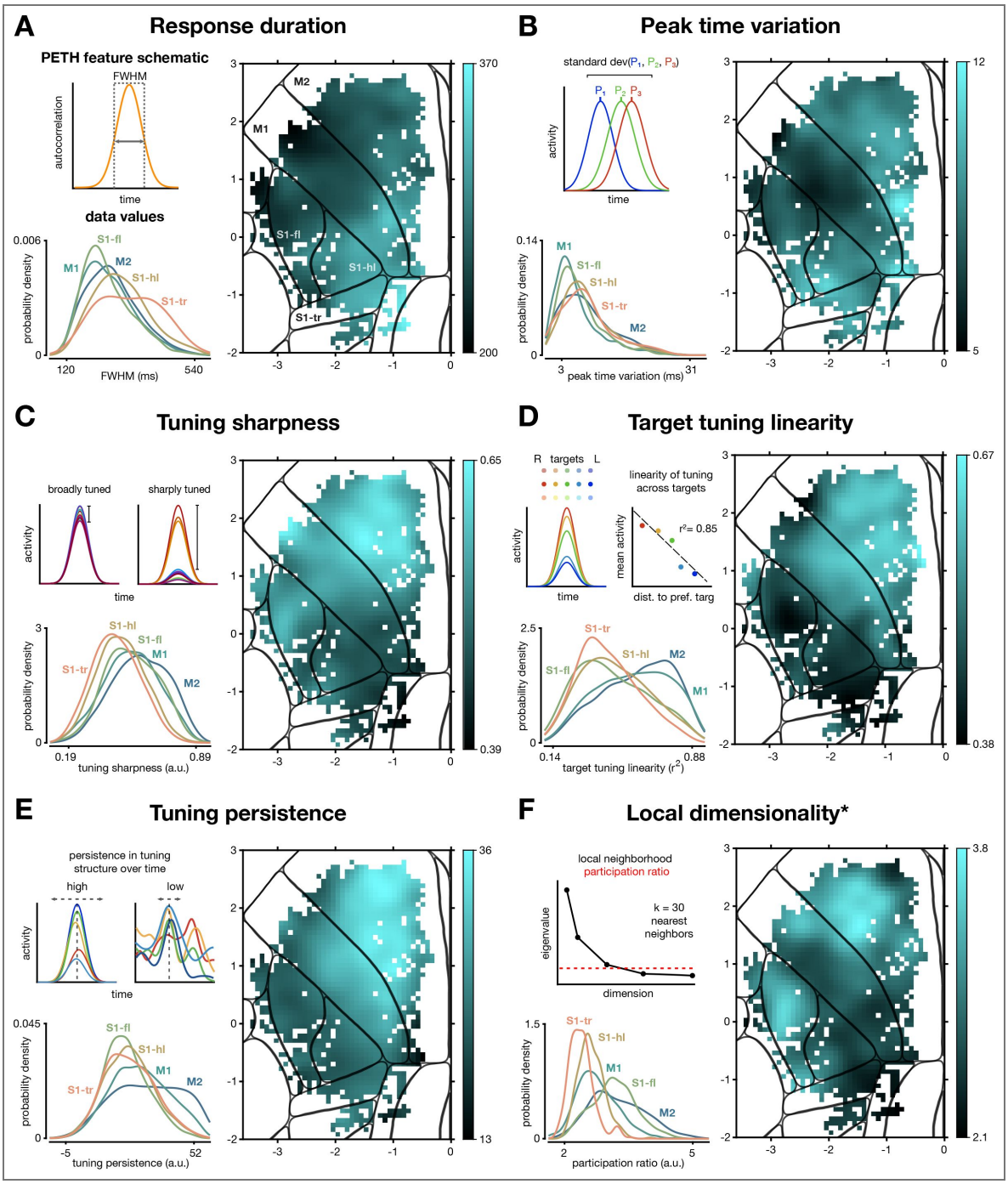


Figure 5. PETH features were organized into distinct spatial patterns in sensorimotor cortex.

(A) *Top left*, schematic for the response duration metric, which measures the autocorrelation width of the trial-averaged trace for each neuron and target. *Bottom left*, histogram of response duration values for all modulated neurons grouped based on anatomical region. *Right*, metric map of the response duration values where bright regions correspond to higher response duration values and darker regions to lower. (B-F) Analogous to A for other metrics. (B) Peak-time variation, computed as the standard deviation of peak times across targets. (C) Tuning sharpness, computed as one minus the ratio of the average peak response (excluding the strongest response condition) normalized by the strongest peak response. (D) Target tuning linearity, the coefficient of determination of the peak response with ordinal distance of the target from the target that evoked the strongest response. (E) Tuning persistence, corresponding to how consistent the tuning was across targets at all lags from the peak (Methods). Only statistically-modulated targets were included for each neuron. (F) "Local dimensionality", the participation ratio of the 30 nearest neighbors to each neuron. Star indicates that this metric was not used for analyses below, because it is not independent for each neuron.

Tuning sharpness

To ask whether a neuron was active for only a few targets or many, we computed a metric we refer to as PETH tuning sharpness (Figure 5C). Neurons attained values near zero if they were similarly active for most targets (i.e., active but broadly tuned), while neurons attained values near unity if they were strongly active for few targets (i.e., sharply tuned). This metric showed that neurons in forelimb-related subregions had sharper tuning than neurons in hindlimb-related regions, with M2 neurons being particularly sharply tuned.

Target tuning linearity

To quantify how linearly a neuron's activity related to target location in physical space, we correlated the 15D vector of mean activity of the neuron for each target with the 15D vector of the targets' ordinal distances from the neuron's preferred target (Methods). The resultant metric map shows that the motor areas were much more linearly tuned with respect to target location than were the somatosensory areas, with M2 even more linear than M1 (Figure 5D).

Tuning persistence

To capture whether a neuron's tuning profile was similar across time points or varied at different time points, we computed a tuning persistence metric. Neurons with high persistence maintained the same tuning across longer stretches of time than neurons with low values. Tuning persistence was highest in the anterior of M2, was moderate in M1, and was low in the sensory regions (Figure 5E).

Local dimensionality

To ask how heterogeneous small neighborhoods of neurons were, we computed a "local dimensionality" metric: for each neuron, we computed the participation ratio (an imperfect but simple eigenspectrum-based measure of dimensionality) of its 30 nearest neighbors. Higher local dimensionality reflects locations where nearby neurons exhibited more heterogeneity in their PETHs with respect to one another in their dominant patterns (but not necessarily the full diversity of signals present in the population if a few large signals are shared). This measure of local dimensionality was highest in a patchy portion of anterior M2, and in anterior S1-fl (possibly corresponding to the tactile region, see Figure 1A from (Alonso et al. (2023)) (Figure 5F). Note that because this metric involved neighborhoods instead of single neurons, it was not included in subsequent analyses.

We also compared our metric maps with maps generated from the top 20 PCs of the PETHs (Methods), rotated using VARIMAX to identify a sparser basis (Musall et al. 2019). These maps also exhibited some spatial structure, but as expected, when compared with our bespoke PETH features created using knowledge of the problem domain, the spatial structure was weaker (Figure 5-figure supplement 1).

PETH features changed more strongly at anatomical and somatotopic boundaries

Each of our PETH feature maps varied across the sensorimotor areas, and appeared to change more abruptly in some locations. We wished to ask whether there were in fact relatively discrete transitions in the PETH feature values across cortical space, or whether these values changed only smoothly. To do so, for each PETH feature we computed the gradient: the vector-valued derivative of the PETH feature with respect to two-dimensional cortical space. We then took the magnitude of this gradient computed at every point in the map (Methods; Figure 6A). Intuitively, this simply asks how quickly the PETH feature is changing at each point on the map. Each map revealed a ridge at the edge of the high-value region (Figure 6A-E), suggesting that the PETH feature maps had edges, albeit soft ones, as opposed to smoothly falling off in value. In the target tuning linearity's gradient map (Figure 6A), this boundary was strongest along the border between M1 and S1-fl, cut through the anterior horn of S1-hl (consistent with projection neuron patterns; Ueno et al. (2018) and Yang et al. (2023)), and separated what is presumably the forelimb

representation from both the hindlimb representation posteromedial and the orofacial representation anterolateral. Other PETH features strongly separated M1 and M2 and somatotopic zones (Figure 6B,D,E). Each of the candidate functional boundaries was present in at least two of the gradient maps. We therefore pooled the identified boundaries by computing an overall gradient magnitude (the quadratic mean of the individual PETH feature maps; Figure 6F). This map makes clear that the PETH features of single cells changed most sharply along both area and somatotopic boundaries. Notably, these boundaries did not correspond with 'edges' in the map of where more or fewer cells were modulated (Figure 3B).

This gradient method has the advantage of producing a smooth map with simple operations. However, to validate the result, we also used a decoding approach. For each of 10,000 randomly-placed candidate boundaries, we attempted to classify which side of the boundary nearby neurons were on using their PETH features (Figure 6G; Methods). We expected that classification would be better when the random boundary happened to match places where PETH features were changing more rapidly. The top-performing 2.5% of random boundaries are plotted in Figure 6H. They aligned well with the boundaries revealed by the gradient method.

Areas exhibited distinct and multimodal activity profiles

From the PETH feature maps and distributions (Figure 5), it was clear that areas exhibited distinct response profiles on average. However, within each area, neurons' responses were heterogeneous. To better understand the distribution of response profiles within and across areas, we used nonlinear dimensionality reduction (t-distributed Stochastic Neighbor Embedding, t-SNE; Hinton and Roweis (2002)). Specifically, we summarized each neuron's response with a five-dimensional feature vector comprising the five PETH features described above, then applied t-SNE to these vectors for the full neural population spanning all five brain areas (Figure 7A). Finally, we found where the neurons from each area were embedded in the t-SNE space (Figure 7B). The resulting two-dimensional t-SNE projection led to three findings. First, single neurons from each region were found across most of the t-SNE space (Figure 7B). This makes clear that much of the diversity of neuron activity patterns was observed in each area. Second, when looking at where neurons in each area most commonly appeared in the t-SNE space, most areas had peaks centered at different locations – though with overlapping distributions (Figure 7C; all pairwise comparisons $p<0.005$, two-dimensional Kolmogorov-Smirnov test). The largest differences were between motor and somatosensory areas, with S1-fl bridging the other S1 subregions and M1. This implies that areas had distinct distributions of their PETH features, but also that there was an interpretable spectrum of typical properties from M2 to M1 to forelimb S1 to hindlimb and trunk S1. Third, the distributions of PETH features within-area were strongly multimodal for several areas. This suggests that the heterogeneity of neurons within each area may form loose response 'types,' especially in M2, M1, and S1-fl. This multimodality motivated further analysis targeted at understanding the range of neural response profiles in each area, and whether neurons with specific response profiles were preferentially found in localized regions of each cortical area.

Neighboring sensorimotor regions shared PETH features in complex spatial patterns

The multimodality of feature vector distributions within-area opens a range of possibilities for how PETH response features could be organized across the cortex, both within and across areas. To work with these multimodal distributions, we fit a generative model to the PETH feature vectors defined above for each area's population of neurons. This approach modeled the feature vectors of an area's neurons as arising from a Gaussian Mixture Model (GMM). Doing so explicitly models different peaks in the distribution as arising from different Gaussian distributions, and retains probabilistic interpretability. In agreement with the t-SNE results, M1, M2, and S1-fl were best modeled by the GMM as having multiple components (Methods): four for M2, three for M1, and two for S1-fl, while S1-hl and S1-tr were best modeled as being unimodal. The Gaussians fitting different modes were well-separated, as expected when fitting a multimodal distribution (Figure 8-figure supplement 1).

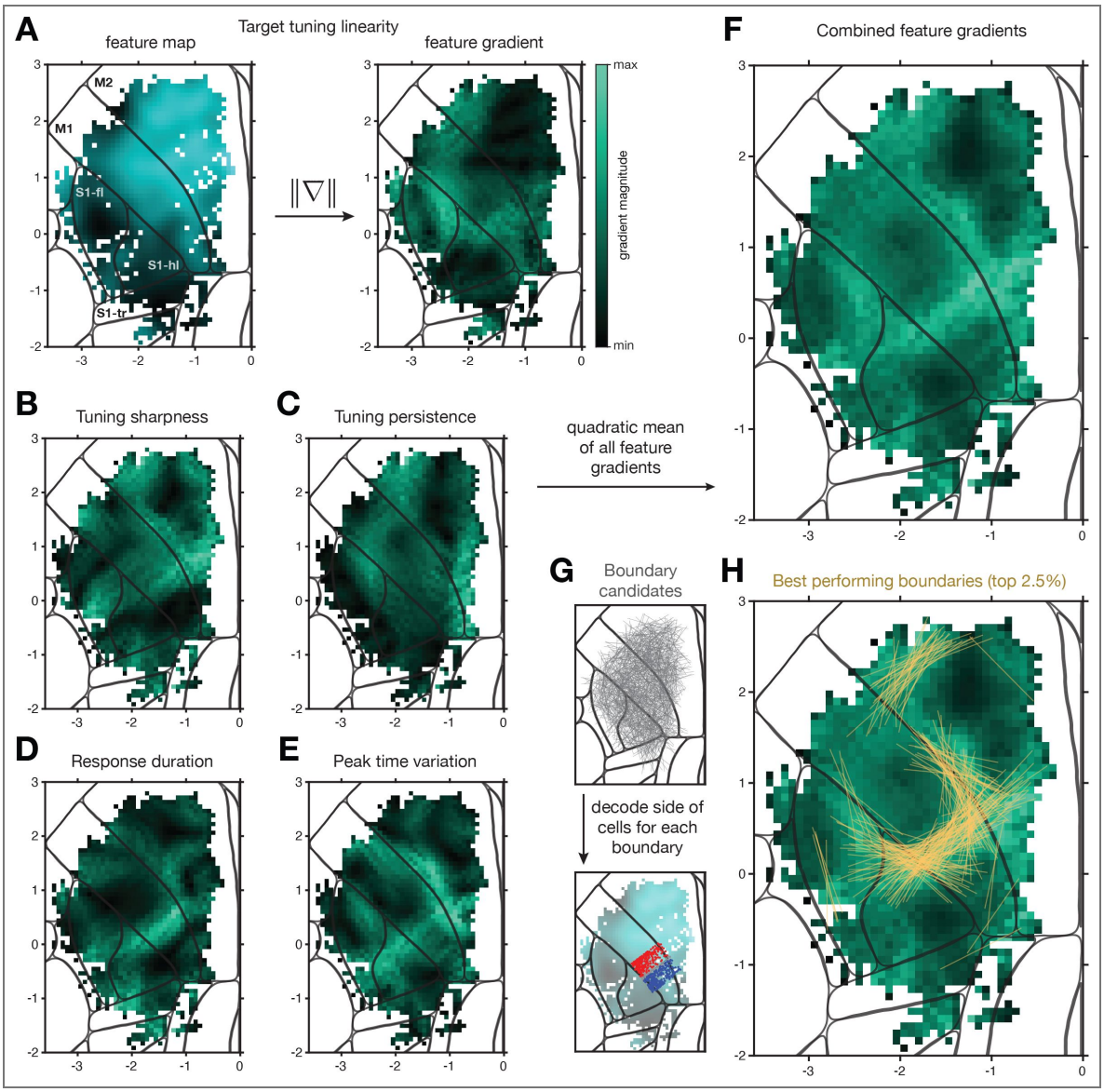


Figure 6. Derivatives of PETH feature maps produced response-property boundaries aligned with anatomy and somatotopy.

(A) Side-by-side comparison of a PETH feature map and its gradient magnitude map. *Right*, bright green values represent large gradient magnitudes and dark values represent small gradients. (B-E) Same as A *right*, for the other PETH features. (F) Quadratic mean over the five PETH features of the gradient magnitude maps. (G) *Top*, 10,000 1-mm random boundaries plotted in gray within the sampled region. *Bottom*, for each random boundary, neurons that were within 500 μm on either side of the boundary were included; neurons within 50 μm of the boundary were excluded. The target tuning linearity feature map is shown underneath for illustration. An SVM was trained to predict which side of the boundary a neuron was on using its PETH feature vector. (H) The top-performing 2.5% of random boundaries (assessed via cross-validation) plotted in yellow on top of the average gradient magnitude map.

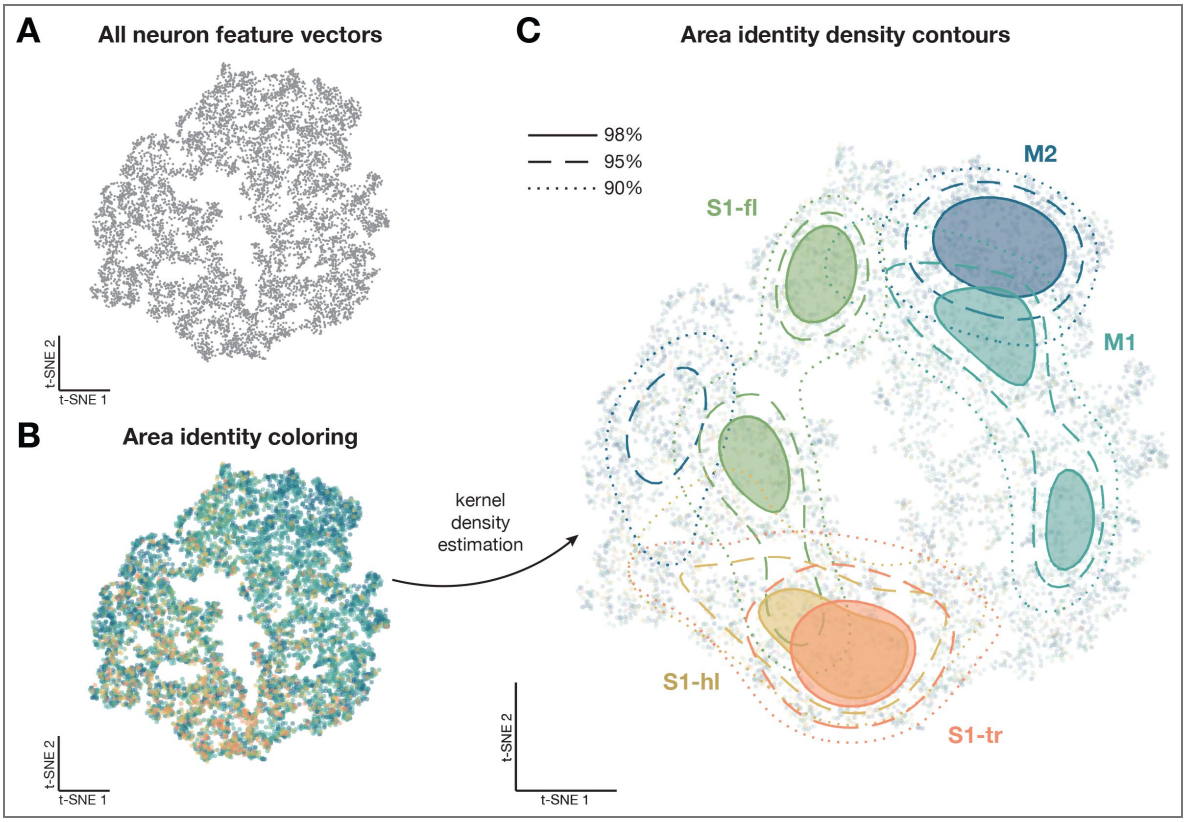


Figure 7. Areas had distinct and multimodal distributions of PETH features.

(A) t-SNE of the five-dimensional PETH features for all neurons. Each gray point corresponds to a neuron. (B) Same embedding as in A, with neurons (points) colored by anatomical area. (C) Contour plots of the embedding from B, using two-dimensional Kernel Density Estimation to produce a smooth density estimate. Three contour levels are shown for each of the five anatomical areas of interest.

We hypothesized that even though the distributions of response profiles were multimodal within-area, different areas might still be well separated. To test this, we computed a data likelihood for each neuron – p(neuron *i*'s PETH features | GMM_{area}) – as arising from each area's distribution (Figure 8A). Intuitively, this produces a measure of how similar each neuron is to the population of neurons in an anatomically-defined area. We then binned and smoothed over neurons by cortical location to produce the maps in Figure 8B-F. This mapping approach is explicitly biased toward finding feature differences between areas, allowing for a direct test of the hypothesis that response profile distributions are area-specific.

When computing this likelihood map using neurons in M2 to define the distribution (Figure 8B), the highest likelihood zone was, as expected, in M2. However, we found that a patch of M1 was strongly similar. If we instead defined the distribution using M1 neurons (Figure 8C), again the highest-likelihood zone was within the defining area. However, two additional features were salient. First, the anterior portion of S1-fl was also strongly similar to M1, in line with previous results (Grier et al. 2026). Second, the posteromedial portion of M1 differed strongly from the rest of M1, consistent with a difference between forelimb and hindlimb somatotopic regions (Maurer et al. 2023). Considering other areas, defining the distribution with S1-fl led to high likelihoods in adjacent M1 and S1-hl (Figure 8D), while defining the distribution with S1-hl or S1-tr produced similar spatial layouts and bled over mostly into the other sensory areas (Figure 8E,F). These results make clear that the working hypothesis – of areas with well-separated feature distributions – is incorrect. Even when building models explicitly from neurons in one area, neurons in portions of some areas were similar to neurons in portions of other areas, and two dissimilar areas could both be similar to the same third location in cortex. These neural populations therefore warranted further dissection.

Subpopulations spatially overlapped and spanned anatomical areas

The likelihood maps above suggested that populations of neurons with similar response profiles were not always confined to the anatomically-defined areas. For example, portions of M1 and S1-fl were strongly similar (Figure 8BC,D). This hinted at a second hypothesis: that a peak in the multimodal distribution from one area might correspond to a peak in the multimodal distribution of a different area. If this were due to a subpopulation that was shared across the areas, we would expect that the M1 GMM would contain a component (a one-area estimate of the subpopulation) with similar characteristics to a component in the S1-fl GMM (Figure 9A). We therefore computed a distance metric (Methods) between each pair of GMM components (Figure 9B). This analysis revealed that many components identified independently in different areas were indeed highly similar, and formed four clusters. We therefore combined components within each of the four clusters (Methods) and took the neurons belonging to each cluster as the subpopulations.

These four subpopulations differed in two major ways. First, by construction, different subpopulations were characterized by distinct PETH features (Figure 9G). Second, their spatial likelihood maps differed strongly as well (Figure 9C-F). One subpopulation, which we refer to as the "Anterior" subpopulation, was almost unique to M2 (Figure 9C). It had uniquely high tuning persistence and tuning sharpness along with highly linear target tuning, arguing for a relatively abstract and external target representation (Galiñanes and Huber 2023). A second subpopulation, the "Forelimb motor" subpopulation, spanned M2 and the forelimb portion of M1 (Figure 9D). These neurons were also strongly linear in their target tuning, but less sharply tuned and with tuning that was somewhat less persistent (i.e., changed more over the course of the movement).

The "Forelimb somatomotor" subpopulation spanned M1 forelimb and orofacial zones and the anterior portion of S1-fl (Figure 9E). This subpopulation was more broadly tuned, was less linearly tuned to target location, changed tuning over the course of the movement (low tuning persistence), and responded very briefly (low response duration). Nevertheless, each neuron in this subpopulation still responded at its own characteristic time in the movement (low peak time variation). Finally, a "Hindlimb somatomotor" subpopulation (Figure 9F) was centered on S1-hl

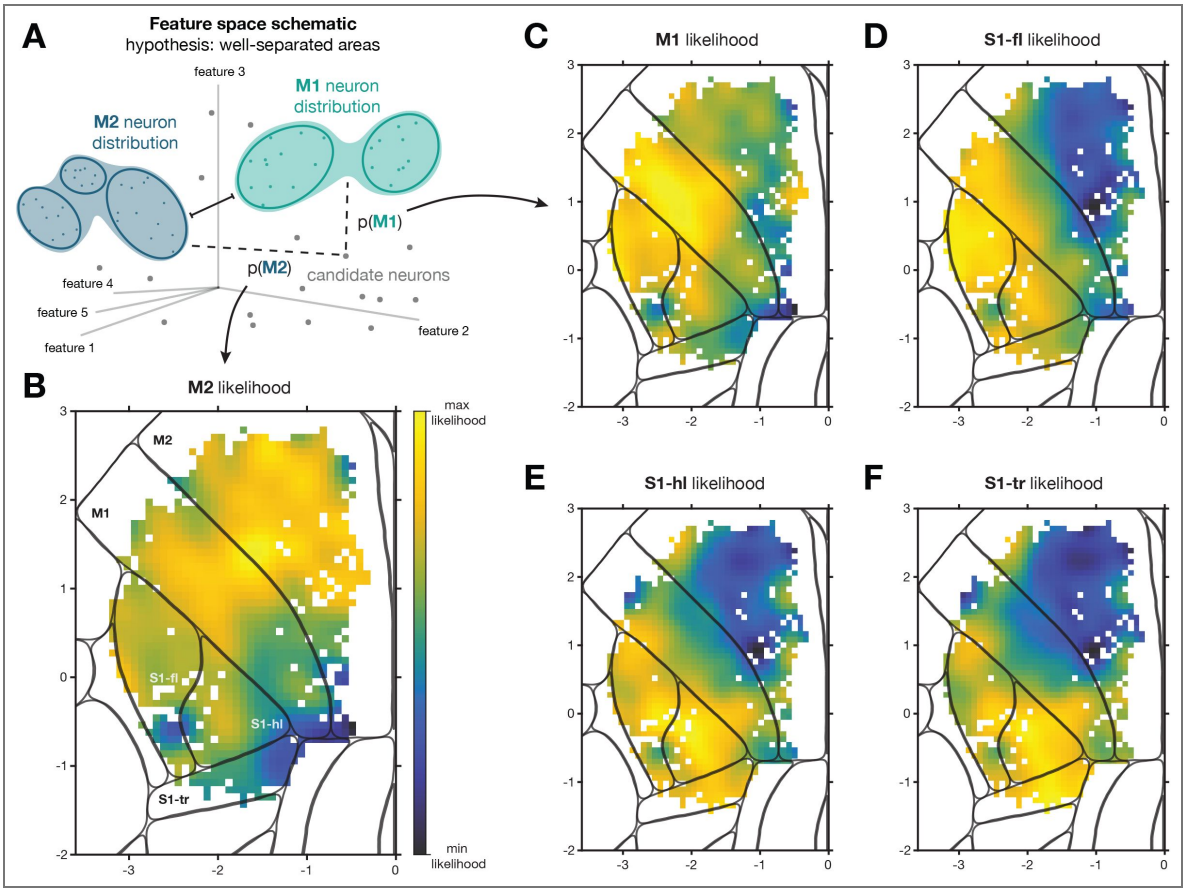


Figure 8. Neural response profiles were shared in patches that spanned anatomically defined regions.

(A) Schematic of analysis with hypothesis that neurons from different areas are well-separated in feature space. A Gaussian Mixture Model (GMM) was fit to the distribution of feature values for neurons belonging to a single anatomically-defined area, then a likelihood was computed for all neurons from all areas using the fit. (B) The resulting likelihood map for M2, where bright yellow pixels correspond to high likelihood regions and dark blue pixels to low likelihood regions. Colorbar scale is logarithmic. The ends of the colormap were set to the maximum and minimum likelihood values for each map. (C-F) Same as B, where the seed regions were M1, S1-fl, S1-hl, and S1-tr, respectively.

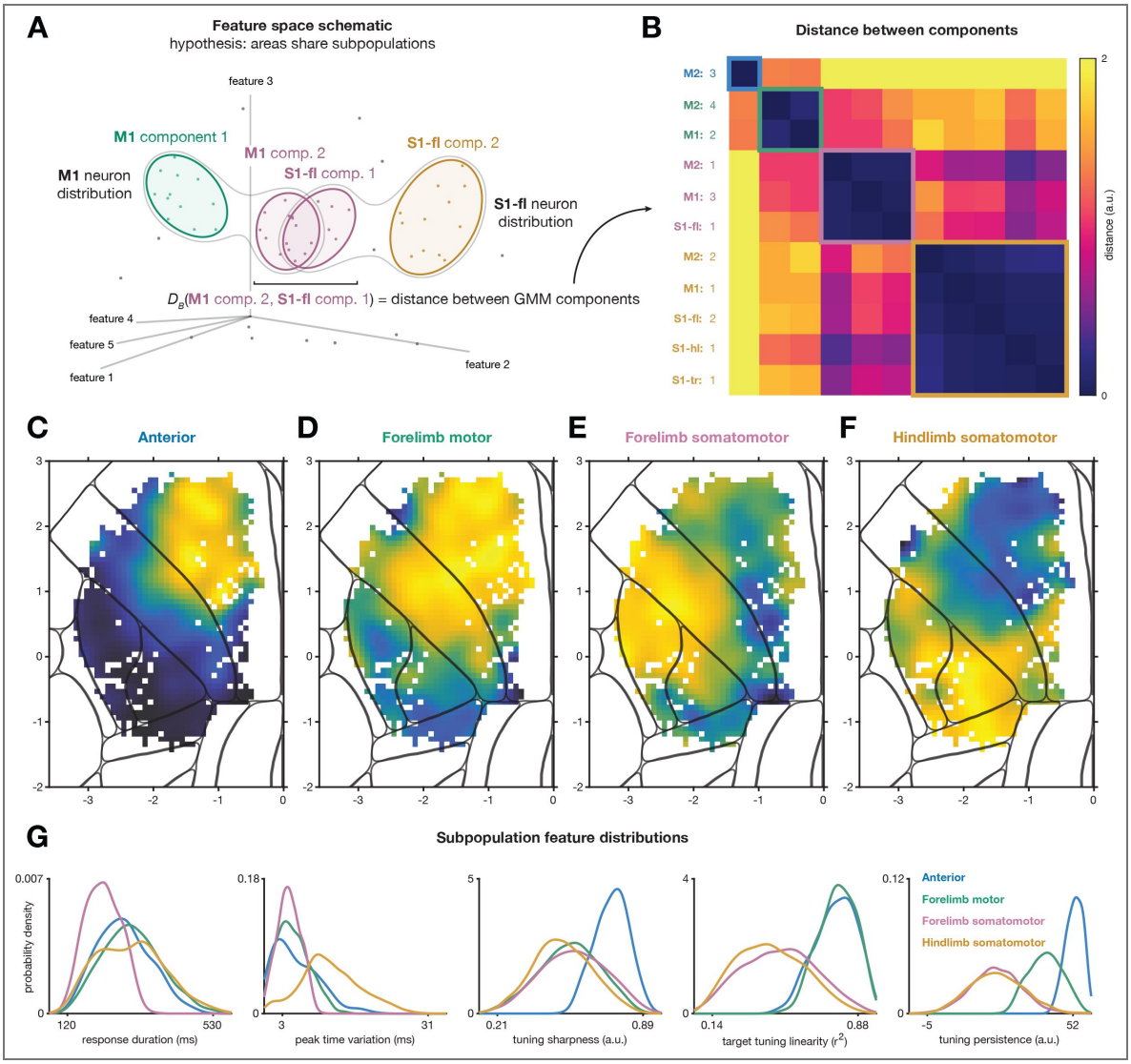


Figure 9. Four subpopulations with distinct response profiles form an overlapping patchwork across sensorimotor cortex.

(A) Schematic of GMM modeling with the hypothesis that components from different areas correspond. Only two GMM components in each area are illustrated for simplicity, with numbers corresponding to those in B. (B) Bhattacharyya distance calculated between all pairs of GMM components across all anatomical regions. The matrix is organized based on cluster identity from hierarchical clustering of the pairwise distances. Each cluster (colored box outlines) is referred to as a subpopulation. Colormap capped to better show any within-cluster detail. Cue-locked data, modulated neurons only. (C) Anterior subpopulation likelihood map. Plotting as in the other likelihood maps with colormap as in Figure 8. (D-F) Same as C but for Forelimb motor (D), Forelimb somatomotor (E), and Hindlimb somatomotor (F) subpopulations. (G) Distributions of the five feature values (subpanels) for each of the four subpopulations (traces)

and spilled into S1-tr, the posterior portion of S1-fl, and the posterior tips of M1 and M2. This subpopulation was similar to the Forelimb somatomotor subpopulation, but each neuron no longer fired at a specific time point in the movement; instead, neurons fired at different times for different targets (high peak time variability). As a qualitative depiction of the response profiles identified with each subpopulation, we plotted the two highest-likelihood cells for each area/subpopulation combination (Figure 9-figure supplement 1). These examples reveal stereotypy in the subpopulation responses across areas, but also show variation across areas, especially for the two somatomotor subpopulations.

To better understand whether the subpopulations were well separated in feature space or not, we examined our clustering further using two approaches. First, for each pair of subpopulations, we trained a logistic regression classifier to separate the 5D feature vectors of the neurons in one subpopulation from the feature vectors of the neurons in another subpopulation, then projected the feature vectors onto this axis (Figure 9-figure supplement 2A). The resulting distributions were well separated along this axis in all but one of the pairwise comparisons (Hindlimb somatomotor vs. Forelimb somatomotor subpopulations). Second, we re-examined our t-SNE of the features, contouring density by subpopulation membership instead of area membership (Figure 9-figure supplement 2B). Despite t-SNE being a quite different nonlinear operation than applying a GMM, subpopulations were again distinct. Interestingly, here the strongest separation was between the Hindlimb somatomotor subpopulation and all other subpopulations. To assess an orthogonal property, we examined the distributions of activity onset times by subpopulation (Figure 9-figure supplement 2C). These distributions appear to be more distinct for the subpopulations than for the areas (compare with Figure 4A). Overall, these analyses support the interpretation that the subpopulation distributions are distinct in feature space.

To confirm that this identification of area-spanning subpopulations was not simply due to our choice of PETH features, we validated these findings by repeating some of our analyses using the top 20 PCs of the PETHs instead of the PETH feature vectors. Using PCA as a preprocessing step in this way is more common in the field. Performing t-SNE on the PCs, analogously to Figure 7, again yielded multimodality (Figure 9-figure supplement 3A). Repeating the analysis of Figure 8 on the PCs produced likelihood maps that were substantially patchier and less distinct between areas, but recapitulated that parts of each area were strongly similar to parts of other areas (Figure 9-figure supplement 3C-G). This indicates that although our choice of PETH features enabled a sharper view of the distinctions between subpopulations, these features were not necessary to detect them.

To determine whether finding distinct subpopulations required recordings from a large area, we directly examined a smaller patch of M1 (480 μm diameter) in an overlap zone (Figure 9-figure supplement 3B, *leftmost*). Applying t-SNE to either our PETH features (Figure 9-figure supplement 3B, *left*) or to the top 20 PCs of the PETHs (Figure 9-figure supplement 3B, *right*) again yielded multimodality. This demonstrates that even with a single ordinary-size field of view, distinct subpopulations in overlapping zones could be readily detected in the context of our behavior.

To validate the clustering itself, we fit a GMM to all neurons without consideration of their area membership. This process produced three components with maps that were nearly identical to three of the four subpopulations identified with the area-aware method (Figure 9-figure supplement 4). The only noticeable difference was that the Anterior subpopulation was lost when all areas were fitted together, likely due to having a lower membership than the other subpopulations. Using this area-agnostic model we also demonstrated that these three main subpopulations were reliably detected in subsets of our data. To do so, we evaluated how similarly two GMMs fit to 50/50 area-wise splits of the data clustered the neurons. The median Adjusted Rand Index (Hubert and Arabie 1985), a measure of clustering similarity, was 0.856 across splits, indicating highly similar clusterings from the data subsets. Together, these controls indicate that the clustering was reliable.

Notably, the high-likelihood regions for different subpopulations spatially overlapped. The Anterior and Forelimb motor subpopulations overlapped throughout M2; the Forelimb motor and the Forelimb somatomotor subpopulations overlapped in the forelimb portion of M1; and the Forelimb somatomotor and Hindlimb somatomotor subpopulations overlapped in a band along the S1-fl and S1-hl border. This indicates that although several of the PETH feature boundaries found in Figure 6 aligned with area borders, several of these boundaries corresponded to the edges of subpopulations that spanned other borders. To confirm this interpretation, we computed the gradients of the likelihood maps directly (Figure 9-figure supplement 5). These maps closely matched those computed from the PETH feature gradients directly (Figure 6), as expected.

Subpopulation members were intermingled and reflected similar response profiles across areas

The overlaps in the subpopulation likelihood maps above imply that members of different subpopulations are spatially intermingled, but it is less clear whether each subpopulation has homogeneous response profiles across space. In particular, the use of likelihoods mixes two properties: the fraction of neurons in a given neighborhood that are members of each subpopulation, and the heterogeneity of response profiles amongst members of that subpopulation. These properties could vary systematically with respect to one another, and the spatial structure shown by the likelihood map does not disentangle them.

We therefore assessed the distribution of subpopulation memberships across cortex independently from response profile variability within subpopulation. To do so, we first classified each neuron into its maximum-likelihood subpopulation, then mapped the subpopulation identity of neurons (Figure 10A). As expected, members of the different subpopulations were locally salt-and-pepper intermingled, including in overlap zones. Also consistent with the likelihood maps, the fraction of each subpopulation in each area differed (Figure 10B). Critically, we found that the prevalence of subpopulation membership across cortex matched the likelihood maps (Figure 10D), implying that the spatial structure in the likelihood maps reflects the spatial variation in prevalence of each subpopulation and not some other structure in the heterogeneity of response profiles.

Finally, we looked for heterogeneity in the PETH features within the footprint spanned by each subpopulation. If the PETH features within a subpopulation varied substantially across areas, a model built on neurons from one area (e.g., M2) would assign lower likelihoods to neurons in other areas (e.g., M1), and vice versa (Figure 10C). To determine if this was the case, we produced likelihood maps for each GMM component of the models fit to each area separately, thresholded these maps to produce a contour, and compared their shapes. Subpopulations had nearly identical spatial footprints no matter which area was used to identify them (Figure 10E). This was true even in cases where only a small fraction of neurons in an area were members of that subpopulation: for example, the Forelimb somatomotor and Hindlimb somatomotor subpopulations could be identified using M2 data, even though there was no high-likelihood zone for these subpopulations within M2 (see Figure 9E-F). Note that this consistency mirrors the low Bhattcharyya distances between corresponding GMM components in Figure 9B, and further validates the use of clustering to combine components from different areas. Together, these results argue that the overlap zones for different subpopulations contained locally intermingled members, and that the PETH features within each subpopulation were similar across the territory covered by that subpopulation.

Discussion

By densely sampling layer 2/3 of mouse sensorimotor cortex during a 15-target reach-to-grasp behavior, we mapped how single cell response characteristics were distributed across five areas. Neurons responded heterogeneously within all imaged areas, and within M2 and M1 in particular, but the average response properties of areas were distinct. In parallel, we identified four subpopulations of neurons with unique activity profiles, which spanned multiple areas and were

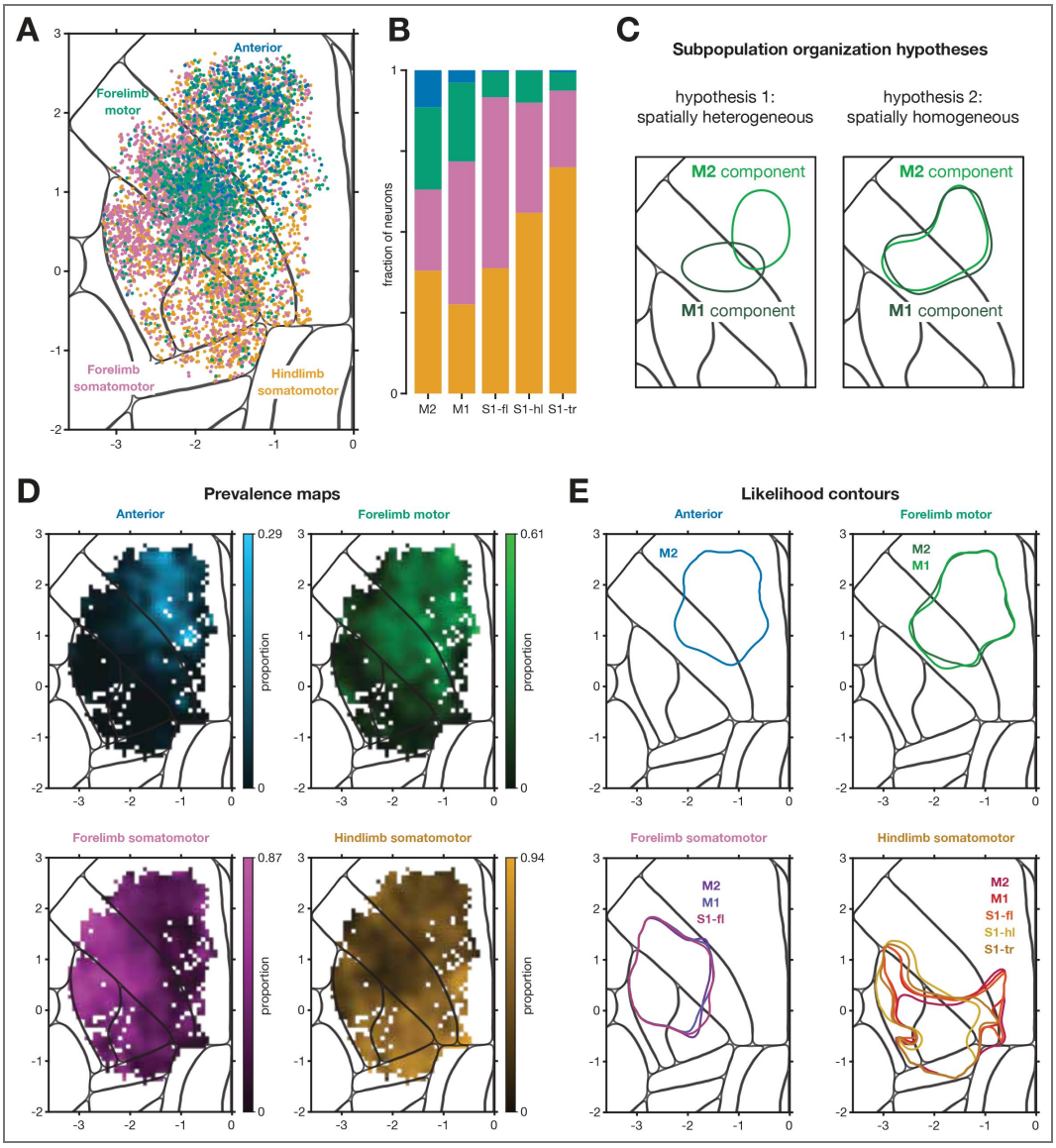


Figure 10. Subpopulations overlapped with their member neurons spatially intermingled.

(A) Neurons were classified into their most likely subpopulation, color coded and plotted on the CCF. (B) Prevalence of subpopulation by area. (C) Schematic for distinguishing two hypotheses, if subpopulations were identified separately using data from each of two areas and the identified high-density zones of neurons were plotted as contours of the likelihood maps. *Left*, expected result if subpopulations were homogeneous in their properties over cortical location; *right*, if subpopulations varied smoothly with cortical location. (D) Maps of prevalence of each subpopulation. (E) Likelihood map contours at 60% level, for each subpopulation identified separately using data from each area containing sufficient subpopulation members.

spatially intermingled in the zones where different subpopulations overlapped. One subpopulation was unique to M2, one spanned M2 and the forelimb portion of M1, one was centered in S1-fl but extended into S1-hl and the forelimb and orofacial portions of M1, and one spanned the likely hindlimb representations of S1, M1, and M2. Members of these subpopulations did not vary substantially with their spatial location, suggesting that the average differences observed between areas were largely, though probably not exclusively, a function of their containing different proportions of the subpopulations. These activity-based maps thus reveal new structure in sensorimotor cortex and complement a variety of existing mapping methods.

The investigation of coding here was somewhat different than the traditional approach. We did not focus on traditional questions of "what" variables are encoded, such as external target location vs. kinematics or reach direction vs. reach duration. Instead, we focused on "how" task-related activity was structured, such as response duration, timing of peaks across targets and tuning linearity. This feature-engineering approach produced a complementary picture to the more common unsupervised approach to feature extraction, using PCA on the PETHs (as used previously; (Esmaeili et al. 2021)). This makes clear that there are potentially important differences not just in the content of the activity (e.g., encoding target vs. movement commands (Grier et al. 2026)), but also its format (e.g., linear encoding vs. nonlinear, persistent vs. brief responses). In particular, we found that anterior areas tended to have responses that were sharply tuned for target and preserved throughout the movement, while posterior areas tended to be broadly tuned for target and exhibited greater temporal heterogeneity.

What, then, might be the functional roles of these subpopulations? While future work relating neural responses to behavior will be essential to answering this question, the present analyses suggest some possibilities. M2 contained neurons that were sharply and linearly tuned for target location, as well as having earlier onsets, longer responses, and consistent tuning over the movement. This might suggest a more abstract (Bernardi et al. 2020) or external encoding (Galiñanes and Huber 2023). S1-hl and S1-tr also had early responses, with activity spanning the movement but differing in complex ways across targets. This may suggest activity related to bracing or stabilizing the animal (Disse et al. 2023; Nandakumar et al. 2021). The responses of S1-fl were briefest and typically weakly tuned for target, but peaked at a tightly-timed moment in the movement for most or all targets, with the peak timing being neuron-specific. These responses may correspond to some specific sensory signal like afferent feedback during paw lift (Grier et al. 2026; Mathis et al. 2017), but understanding what that signal would be will require future study. The shape of the Forelimb motor subpopulation is similar to the shape of a recently proposed organization based on corticofugal projection patterns, the Medial Anterior Cortex (MAC) (Yang et al. 2023). This would be consistent with MAC's proposed role in reaching. Finally, M1 was at the center of the overlap between subpopulations. This is reminiscent of its suggested role in taking top-down commands from M2 and elsewhere and bottom-up sensory input from S1 to actively control the limb (Galiñanes and Huber 2023; Heindorf et al. 2018; Terada et al. 2022). In each of these cases, determining the relationships of the observed activity patterns to function will require specific attempts to link the activity to kinematics, target location, sensory feedback, and more; these relationships will be addressed in future work.

The area-spanning nature of the subpopulations could arise from a number of possible organizations of neural circuits. One hypothesis is that each subpopulation represents a single recurrent circuit that is distributed across areas. In this hypothesis multiple circuits exist in parallel within each area, each strongly connected within the subpopulation and less strongly connected with the other subpopulations. An alternative hypothesis is that a single area contains the circuit that generates the signals characteristic of a single subpopulation, then this activity is broadcast to other areas through the strong connectivity of sensorimotor cortex (Muñoz-Castañeda et al. 2021; Zingg et al. 2014). This broadcast would then cause that characteristic activity pattern to appear in the recipient areas. Recording neurons that project from one area to another, and recording from other layers corresponding to later stages of processing and outputs, will be critical to distinguishing these hypotheses and more fully understanding these circuits.

The maps produced here, using response properties of neurons during task performance, aligned with anatomical and stimulation-based maps in some respects but differed in others. We identified response-property boundaries that aligned closely with the anatomical borders between areas M2, M1, and S1. We also identified response-property boundaries consistent with somatotopy, separating forelimb from hindlimb and forelimb from orofacial representations. Interestingly, these somatotopic boundaries were generally stronger than those corresponding to "true" area borders, despite the known coarseness of somatotopic maps in primate motor areas (Mitz and Wise 1987; Neafsey et al. 1983; Qi et al. 2010; Van Acker et al. 2013); whether the maps are similarly coarse in mouse remains to be tested. Some of our maps also distinguished the tactile portion of S1-fl from the remainder of the area, and grouped the proprioception-receiving portions of S1-fl and M1 (Alonso et al. 2023). The spatial distribution of modulated cells in Figure 3 suggests a distinction between the caudal forelimb area (CFA, involving M1 and S1-fl) and the rostral forelimb area (RFA) in M2, while the feature gradient boundaries suggest a distinction between M1 and M2 more generally. The absence of a clearly delineated RFA was surprising, given its distinct projection patterns (Carmona et al. 2024; Hira et al. 2013b; Wang et al. 2018) and functional differences from CFA (Kristl et al. 2025; Morandell and Huber 2017; Saiki-Ishikawa et al. 2025), but our results might suggest that the activity in layer 2/3 of RFA does not differ markedly from other nearby subregions of M2. The difference in spatial footprints of the Forelimb motor and Hindlimb somatomotor subpopulations also loosely corresponded to the spatial distribution of projections to the lateral rostral medulla (denoted the lateral anterior cortex or LAC in (Yang et al. 2023)), but assessing exact correspondence will require methods not employed here. Finally, our approach did not clearly identify the anterolateral motor area (ALM; (Chen et al. 2017; Inagaki et al. 2018; Li et al. 2015)) within M2, though this may be a consequence of the time period analyzed; a time window targeted to the period of withdrawal and water consumption might more clearly delineate this region. These various anatomical distinctions might appear in activity patterns during other tasks that require different coordination between subareas, with analysis of other time periods, with use of additional response features, or with recordings from deeper layers. Identifying such functional differences is left to future work.

Crafting additional PETH features, or using end-to-end neural network approaches to discover other features, might enable the discovery of additional structure (Minderer et al. 2019; Wang et al. 2023b). For example, our PETH features were chosen to be invariant to the onset time of activity, but these onset times were markedly later in lateral M1 than in adjacent M2 or S1-fl. Including onset times, using a wider window of time that includes more of the reward/licking period, aligning data to other behavioral events, or adding other PETH features would presumably result in finer subdivisions of sensorimotor cortex. This suggests that although our current methods find a large amount of structure in sensorimotor cortex, more structure in the responses surely exists. Relatedly, using other tasks might reveal additional functional differences, or minimize ones we identified. Similar explorations of data from other behaviors is likely to shed additional light on the organization of responses in these areas.

Our findings of larger-scale structure, in the form of area-spanning subpopulations, may intersect with three prior findings. First, several widefield calcium imaging studies have also found that S1-fl activity groups with activity in portions of M1 and S1-hl (Costa et al. 2025; Musall et al. 2019; Saxena et al. 2020; Zatka-Haas et al. 2021). Relatedly, orofacial representations share similar relationships with behavior across M2, M1, and S1 (An et al. 2025; Musall et al. 2019); and both these spatial patterns of coactivation reflect anatomical connectivity (Zingg et al. 2014). These previous findings of larger-scale grouping are consistent with our results. Second, in several cortical areas, populations with different projection targets have known functional differences (Chen et al. 2013; Currie et al. 2022; Economo et al. 2018; Hwang et al. 2019; Kim et al. 2018; Li et al. 2015; Park et al. 2022). However, these specific populations reside mainly in layer 5, while imaging here was performed more superficially. Third, the existence of subpopulations of neurons in sensorimotor cortices have been found independently by using genetic identification of excitatory neuron subtypes, and these subpopulations exhibit distinct activity patterns (Hira et al. 2013a; Li et al. 2024; Mohan et al. 2023; Musall et al. 2023). Whether those subpopulations correspond to the ones found here cannot yet be reconciled.

However, many of the genetically-identified neuron subtypes in those studies are present only in deeper cortical layers, while the imaging performed here was exclusively superficial. This suggests that the subpopulations we have identified are unlikely to correspond closely with the particular subtypes found thus far via genetic tools.

The distributions of PETH properties we observed were multimodal, but these distributions were continuous; we did not observe tight clusters in feature space separated by gaps. In interpreting the observed multimodal distribution, we note that a uniform distribution can be transformed into a multimodal one, or vice versa, by a continuous nonlinear function. This is a fundamental problem whenever searching for structure in the distribution of neural activity: even examining the raw spike rate of a neuron involves the spike threshold nonlinearity, and what property counts as the 'fundamental, linear' one is almost a philosophical question. Nonetheless, the potential for nonlinear measurement makes it challenging to determine whether there are true "categories" of neurons (Kaufman et al. 2022; Posani et al. 2025). If the interpretation of separate modes as subpopulations were incorrect under some relevant nonlinear transformation, our results would nevertheless imply that M2 has the highest diversity of responses, with shrinking variety toward the posterior as well as changes in the average response profile. If this alternate interpretation were the strongest conclusion possible, it would still be essential to understanding the layout of function across sensorimotor cortex. Confirming that these modes are more deeply meaningful will require future external validation.

Several limitations of the present work should be noted. First, the data here include only superficial layer 2/3. Activity patterns in other layers, or even in the deeper portion of layer 2/3, may differ (Currie et al. 2022; Heindorf et al. 2018; Masamizu et al. 2014). Second, while we took great care in aligning our FOVs to the Allen CCF, we did not analyze the anatomy of fixed tissue for each mouse. Limitations in our alignment, differences between the anatomy of the strain of mice used here from the strain used for building the atlas, and possible mouse-to-mouse variation in the borders of areas all introduce jitter. The sharpness of the boundaries we identified therefore serves as a lower bound. Third, although we achieved coverage of a large swath of sensorimotor cortex, we sampled motor areas more densely than sensory and not every area was densely sampled in every mouse. We therefore may have underestimated heterogeneity in S1, which was sampled less. Fourth, although we obtained high-speed video of the behavior, this work does not consider the kinematics. The relationship of these neural data to the detailed kinematics will be treated at length in future work. Fifth, this work considered only trial-averaged responses, not single trials. This choice enabled us to compare much larger areas of cortex by pooling across sessions, but precluded many types of analysis including of noise correlations (Kiani et al. 2015; Kohn et al. 2016; Zohary et al. 1994). Future work could leverage simultaneous recording of multiple areas to investigate single-trial co-fluctuations in neural activity within subpopulations and across areas.

In summary, these findings provide a response-based atlas of sensorimotor cortex and reveal overlapping subpopulations with intermingled members. The nature of the circuits that underlie these subpopulations, and the content of communication between areas more broadly, remain questions of fundamental importance for understanding how movement is generated and feedback incorporated.

Methods

Subjects and surgical procedures

All procedures were approved by the University of Chicago Institutional Animal Care and Use Committee. For two-photon imaging, six male mice aged 8-12 weeks were used. Mice were crosses of Slc17a7-IRES2-Cre-D (JAX strain 037512; Harris et al. (2014)) and TIGRE2-jGCaMP8s-IRES-tTA2 (JAX strain 037719; (Sweeney et al. 2025)) which express GCaMP8s in most glutamatergic neurons (Wang et al. 2023a). Each animal underwent a single surgery, then were individually housed on a reverse 12-h light/dark cycle, with an ambient temperature of 22 °C and a humidity of 58%. Experiments were conducted in the afternoon, during the animal's dark cycle.

Anesthesia and basic surgical preparation were as described in (Grier et al. 2026). A circular craniotomy was outlined with a 5-mm biopsy punch and completed with a dental drill, centered over the left forelimb M1 at approximately 0.4 mm anterior and 1.5 mm lateral of bregma. The craniotomy was cleaned with SurgiFoam (Ethicon) soaked in phosphate-buffered solution (PBS). If the dura remained intact, a durotomy was performed with a needle.

A retrograde virus carrying a tdTomato construct (AAVretro-tdTomato, diluted to 1×10^{13} particles per mL in PBS, Addgene stock 59462-AAVrg) was injected in S1-fl (mice 1-3; 0.0 mm A, 2.65 mm L relative to Bregma) or the center of the forelimb representation of M1 (mice 4-6; 0.4 mm A, 1.5 mm L). 100 nL of virus was pressure injected (NanoJect III, Drummond Scientific) at a depth of 300 μm from the cortical surface over 5 min. The injection pipette was kept in place for another 3-6 minutes to ensure viral dispersion before being removed slowly over 2-3 minutes. This injection retrogradely labeled cell bodies in other areas that sent projections to the injected area. In this paper the labeling was used solely for stabilizing the imaging plane (see below) throughout the cortex.

The craniotomy was sealed with a 5 mm #1 round cover glass (Thomas Scientific) glued in place first with tissue glue (VetBond, 3M) and then with cyanoacrylate glue (Krazy Glue) mixed with black dental acrylic powder (Ortho Jet; Lang Dental). In all mice two layers of MetaBond (Parkell) were applied, and then a custom water jet-cut stainless steel head bar was affixed to the skull with black dental acrylic. Animals were awoken by administering atipamezole via intraperitoneal injection and allowed to recover at least 3 days before water restriction.

Behavior and training

The behavioral task (Figure 2A) was a variant of the multi-target water-reaching task of Galiñanes and Huber (2023). Water-restricted mice were head-fixed with their forepaws on paw rests (bent partially threaded screws) and the hindpaws and body supported by a custom 3D-printed translucent seat. Trials were triggered after the animal held both paw rests for 100 ms. After an additional 100 ms hold period a water spout (22 gauge, modified 90-degree bent, 1-in blunt dispensing needle, McMaster) was moved into one of 15 positions in front of the mouse's face via a set of three motors (LSM050B-T4A, Zaber). The spout was customized, with ~2 mm of the side facing away from the animal ground down; this caused the water to cling to the last few mm of the spout and not just hang from the tip. The animals were then required to hold the paw rests continuously for an additional 400-600 ms delay period to trigger the delivery of a water droplet on the spout. If the animals released the paw rests briefly during this time the delay was reset. If the paw rest release exceeded 1 second contiguously the trial was aborted and a 10 second penalty period occurred. Brief releases occurred at variable rates across animals (20-85% of trials) and delay-aborted trials rarely occurred (<1% of trials). Animals were allowed a maximum of 2000 ms to achieve paw rest contact for the entire delay; if this did not occur the spout was removed to the neutral location and a 5000 ms time-out was triggered (<3% of all trials across animals). Once the paw rests were held continuously for the variable delay time, a 4000 Hz tone was played by stereo speakers as the Go cue and a 2-3-μl droplet of water appeared on the spout. The cue tone lasted 500 ms or until the mouse made contact with the water spout, whichever came first. The mouse could grab the water droplet and bring it to its mouth to drink anytime after the tone began. The spout was removed 4000 ms after first contact, triggering a variable ITI of 3000-7000 ms, and automatically wiped against a small sponge to remove any remaining water. Both the paw rests and spouts were wired with capacitive touch sensors (Teensy 3.2, PJRC). Control software was custom-written in MATLAB R2018a using PsychToolbox 3.0.14. Touch event monitoring and task control were performed at 60 Hz using the Teensy.

The mice were trained to make all reaches with the right paw and to keep the left paw on the paw rest during reaching. For the first 1-3 days, mice were head-fixed and provided water directly to their mouths to acclimate them to head-fixation and reward collection. During this period, the animals were encouraged to make consistent contact with both paw rests to trigger water delivery to their mouths. After acclimation and consistent trial initiation, the spout was withdrawn to a position on the right side of their face intended to elicit forelimb reaching with the right arm. If

animals were reluctant to reach, small droplets of water were placed on their right whisker pad with a blunt syringe to induce grooming and water collection. After animals reliably performed reaches toward the rightward spout location (80-90% success), its position was moved medially and anteriorly in gradual increments within a single session until it aligned with the midline. Once mice could reliably reach to the central spout position, two new target locations were introduced on either side (approximately 1 mm from center). Once in this three-target training stage, each target was cued in blocks of 50 trials. The size of the blocks was reduced as the mice progressed in their training until targets were interleaved randomly. After 80-90% success at the three target configuration, two additional targets were added further lateral on the left and right sides, creating a five target semicircle. The targets were then progressively spaced apart along the semicircle until they reached their final azimuthal spacings of ~1.72 target-to-target and ~6 mm total. Variation in the Z position was introduced by adding an additional row of targets above and below the existing row and gradually increasing the inter-row distance until the final Z spacing of 1 mm was achieved. This training procedure took 10-20 days. Data presented here were collected after 3-6 weeks' experience with the task.

Two-photon calcium imaging

Calcium imaging procedures were similar to those described previously (Grier et al. 2026). Briefly, imaging was performed with a Neurolabware 2p microscope running Scanbox 4.1 and a pulsed Ti:sapphire laser (Vision II, Coherent). Depth stability of the imaging plane was maintained using a custom plugin that made automatic movements of the objective up to every ~10s based on comparing red-channel data with an initially-acquired image stack. Offline, images were run through Suite2p to perform motion correction, region-of-interest (ROI) detection, and fluorescence extraction. ROIs were manually curated using the Suite2p GUI. Fluorescence was neuropil-corrected using a coefficient optimized by minimizing the mutual information between the corrected trace and the neuropil trace (Allen 2025). As previously described (Grier et al. 2026; Zhu et al. 2022), fluorescence was then detrended and passed through OASIS using the 'thresholded' method, AR1 event model, and limiting the tau parameter to be between 300 and 800 ms. To put events on a more useful scaling, for each ROI, we found the distribution of event sizes, smoothed the distribution (ksdensity in MATLAB, with an Epanechnikov kernel and log transform), found the peak of the smoothed distribution, and divided all event sizes by this value. This rescales the peak of the event size distribution to have a value of unity.

As previously described, when time-locking the data for visualization, modulation testing, or further analyses, the deconvolved events were resampled into 10-ms bins. This resampling was performed by assigning a fraction of each event into the new 10-ms bins proportionally to how much the 10-ms bins overlapped the original ~32.2 ms frames, and taking into account the exact time each ROI centroid was sampled based on its position within the field of view (FOV).

All cells were imaged between 140 and 210 μm below the cortical surface, in superficial layer 2/3. We chose this depth range both because it yielded large numbers of modulated neurons, and because GCaMP expression in these animals was lower in deeper layer 2/3 of S1. Imaging in this limited depth range therefore produced more comparable data across areas.

Measuring modulation

To determine whether a neuron was modulated in the task, we used an adapted ZETA method (Montijn et al. 2021) as described in (Grier et al. 2026). Briefly, trial-averaged responses were compared to shuffled trial-averages computed by circularly permuting each trial by a random offset. Modulation was computed as the maximum difference between the real and shuffled data. This procedure was performed separately for each target and for locking data to the cue (−300 ms to +300 ms), lift (−200 ms to +400 ms), and first contact (−300 ms to 300 ms). For each locking event (cue, lift, and first contact), neurons with any p-values ≤ 0.01 were considered modulated. For all analyses, only neurons modulated by the relevant locking event were included. Note that this measure looks for modulation over time to any target; it is indifferent to whether the neuron exhibits tuning across targets.

Mapping cell positions onto a common coordinate frame (Allen atlas)

Each 2P FOV was aligned based on the vasculature to an image of the entire window captured via widefield imaging. S1-fl was identified by paw vibration (Alonso et al. 2023), with imaging and preprocessing performed exactly as described in (Grier et al. 2026) except that the camera was upgraded to a pco.edge 5.5 CLHS sCMOS camera (Excelitas). The vibration map was aligned to the widefield vasculature image via shared vasculature landmarks. This composite image was then scaled and translated to best align the S1-fl activation shape to its CCF boundary, while aligning the central sinus vasculature to the midline overlapping bregma. See Figure 1-figure supplement 1 for vibration map to Allen CCF alignment across animals. 2P images from the cortical surface from each session were then scaled to 720 μm width by 980 μm height and aligned to the registered vasculature via rigid transformations. This series of transformations produced global coordinates for each ROI from its location within the 2P FOV.

Summarizing subregions from prior literature

To produce the map in Figure 1A, outlines of subregions described by prior literature were extracted. Our review of the literature is non-exhaustive and serves simply to illustrate the heterogeneity of function that has been described in previous studies. Corresponding figures were overlaid with the Allen CCF and rigidly transformed to qualitatively match anatomical boundaries as closely as possible. Outlines were then extracted by tracing the subregion boundaries by hand.

From anterior to posterior we describe the acronyms and the corresponding citations as follows: Secondary motor cortex (M2), or MOs in the Allen CCF. Anterior lateral motor area (ALM), as originally described in Komiyama et al. (2010) and further delineated in Guo et al. (Guo et al. 2015) and Chen et al. (2017). Rostral Forelimb Area (RFA) as described in Tennant et al. (2011) and further delineated in Carmona et al. (2024). Rostral Forelimb Orofacial area (RFO) as described in An et al. (An et al. 2025), and referred to as tjM1 in Mayrhofer et al. (2019). Secondary vibrissal motor area (M2-vb) as described in Mayrhofer et al 2019 as the frontal region activated by C2 vibrissal stimulation, and called medial motor area or MM in Chen et al. (2017). Primary motor cortex (M1) or MOp in Allen CCF. Primary forelimb motor cortex (M1-fl) as described in Munoz-Casteneda et al. (2021) via tracing and characterized in Grier et al (2026). Primary hindlimb motor area (M1-hl) as traced from hindlimb musculature by Maurer et al. (2023), and sensory cortical areas by Zingg et al. (2014). Primary forelimb somatosensory area (S1-fl), or SSp-ul in Allen CCF. Primary proprioceptive and tactile forelimb somatosensory areas (S1-pro and S1-tac) as originally identified in Alonso et al. (2023). Primary hindlimb somatosensory area (S1-hl), or SSp-ll in Allen CCF. Primary trunk somatosensory area (S1-tr), or SSp-tr in Allen CCF. Primary mouth somatosensory area (S1-m), or SSp-m in Allen CCF. Primary nose somatosensory area (S1-n), or SSp-n in Allen CCF. Primary vibrissal somatosensory area (S1-bfd) or SSp-bfd in Allen CCF. Anterior visual cortex (A) or VISa in Allen CCF. Retrosplenial cortex (RSC), or RSPagl, RSPd and RSPv in Allen CCF. Cortical cervical spinal cord (CSC) and lateral rostral medulla (latRM) projections as traced in Yang et al. (2023).

Computing map images from single-cell properties

To quantify and visualize the spatial distribution of properties of imaged neurons, we used a binning procedure to prevent biases due to uneven sampling. We binned the neurons into 80 μm-wide square pixels, then averaged the property of interest over all neurons within the pixel. Pixels that contained zero neurons were excluded from further analysis. The images were then Gaussian smoothed (s.d. 150 μm). To Gaussian-smooth the images while excluding empty pixels and accommodating irregular map edges, the final normalization for each pixel was the sum of the weights of contributing pixels, excluding empty pixels. For modulation images, each neuron was assigned a zero if not statistically modulated or a one if modulated. For PETH feature maps, each neuron was assigned its PETH feature value. The one-dimensional pixel values were then mapped onto a perceptually-linear colormap (Kovesi 2015).

PETH feature computations

Response duration

We quantified the response duration of single neurons by analyzing the autocorrelation structure of their responses over time. For each neuron, we first computed the autocorrelation function $r(\tau)$ of the trial-averaged response over time $x(t)$ to each target separately:

$$r(\tau) = \frac{\sum_t (x(t)-\bar{x})(x(t+\tau)-\bar{x})}{\sum_t (x(t)-\bar{x})^2}$$

where τ represents time lag, and $\bar{x}$ is the mean of the target-specific response over time. Only targets for which the neuron was statistically significantly modulated were included. From this autocorrelation, we determined the full width at half maximum (FWHM). The mean FWHM was calculated across all the targets for which the neuron was modulated (up to 15), resulting in a scalar measure of response duration for each neuron. Higher values of mean FWHM indicate neuronal responses that are more sustained (with a broader peak), whereas lower values indicate a more transient response.

Peak time variation

Some neurons exhibited firing rate peaks that were very similarly timed for all targets, while other neurons did not. To quantify this, each neuron's PETH was first computed as described above, then for each condition (target) c, the peak response time t_c^{peak} was identified as the time bin at which the trial-averaged trace reached its maximum amplitude. To prevent the inclusion of conditions with minimal modulation, we restricted this analysis to conditions whose peak amplitudes were at least 25% of the maximum peak amplitude across all 15 conditions for each neuron. The variability in peak timing was then quantified as the standard deviation of the selected peak times:

$$\sigma_{\text{peak}} = \sqrt{\frac{1}{n}\sum_{c=1}^{n}\left(t_c^{\text{peak}} - \bar{t}^{\text{peak}}\right)^2}$$

where n is the number of conditions meeting our inclusion criterion and $\bar{t}^{\text{peak}}$ is the mean peak time across conditions above the threshold. A higher σ_{peak} indicates greater variability in the timing of peak responses across targets. To measure the signal-to-noise ratio for each neuron, which we used to characterize how strongly the peak time variation metric related to response consistency, we used a simple proxy from previous work (Kaufman et al. 2016). Specifically, for each neuron, we took the maximum firing rate for any time and condition minus the minimum firing rate for any time and condition, and divided this difference by the maximum s.e.m. for any time and condition.

Tuning sharpness

Tuning sharpness measures how selective each neuron is to some targets over others. For each neuron, we obtained the peak response $v(c)$ for each condition (i.e., target) c. The tuning sharpness then computes the average response as a fraction of the max response, and subtracts this from one

$$\text{Tuning sharpness } = 1 - \frac{1}{14}\sum_{c \neq c_{\max}} \frac{v(c)}{v(c_{\max})}$$

This metric ranges from near one, meaning a neuron is highly selective for a single condition, to near zero, indicating a neuron achieves a similar peak firing rate for all targets.

Target tuning linearity

To assess how linearly related a neuron's response profile was to the target locations in physical space, we performed an analysis related to Representational Similarity Analysis (RSA; Kriegeskorte et al. (2008)) but limited it to linearity with respect to an 'anchor' target. To do so, we first computed a matrix containing the ordinal distances between all pairs of targets, and a matrix of the 'distances' between the firing rates for each pair of target responses (averaged over the time window from 200 ms prior to the cue to 400 ms post-cue). We then computed the coefficient of determination (R^2) between row m of the target distance matrix and the corresponding row m of the neural distance matrix, for all rows m. Finally, we took as the target tuning linearity the

highest of the 15 R^2 values. Intuitively, this looks for a linear relationship between firing rate and target distance with respect to an 'anchor' target that maximizes the correlation. Values near one indicate a neuron whose firing scaled perfectly linearly with target location assessed in this way.

Tuning persistence

For each neuron, we first computed the PETH as described above. Then, for each neuron *i*, we identified the target for which the neuron achieved the greatest peak response, then found the time point t_{rmax} of that peak. An anchor tuning vector $\mathbf{v}_i(t_{rmax})$ was computed, whose entries are the PETH amplitudes across targets at time t_{rmax}. At every time point *t*, we measured the similarity of the instantaneous tuning profile vector $\mathbf{v}_i(t)$ to the anchor vector by the Pearson correlation

$$r_i(t) = \mathrm{corr}\left(\mathbf{v}_i(t), \mathbf{v}_i\left(t_{\mathrm{rmax}}\right)\right)$$

This yields a "tuning-distance" curve $r_i(t)$ that falls off as the neuron's PETH tuning profile deviates from that at $t = t_{rmax}$. To summarize these correlation plots into a single value, we computed the area under the tuning-distance curve

$$\mathrm{AUC}_i = \int r_i(t)\mathrm{d}t$$

using trapezoidal integration over the sampled timepoints. Neurons with slowly decaying correlations (large AUC) achieved higher tuning persistence, whereas rapidly decaying neurons (small AUC) received low tuning persistence.

Local dimensionality

We estimated the effective dimensionality of local population activity by computing a common and simple metric, the participation ratio (PR), in the spatial neighborhood of each recorded neuron. This metric is limited to only considering the singular value spectrum of the local activity, but when applied exclusively to modulated neurons it provides at least a coarse idea of the variety of responses. For each neuron, we identified its $K = 30$ nearest neighbors (including the anchor neuron) by Euclidean distance in anatomical (*x*, *y*) coordinates. For each neighbor, the trial-averaged PETH was vectorized; stacking these vectors formed a data matrix *X* whose columns correspond to neurons. We then computed the eigenvalues λof the covariance matrix of *X*. Finally, the participation ratio for that neighborhood was taken as:

$$\mathrm{PR} = \frac{\left(\sum_{j=1}^{K}\lambda_j\right)^2}{\sum_{j=1}^{K}\lambda_j^2}$$

PR increases as variance is distributed more evenly across orthogonal components (higher local dimensionality), approaching 1 when variability is dominated by a single component (neuron) and approaching *K* (number of neighbors) when contributions are comparable across all components. The procedure was repeated with each neuron serving as the anchor, yielding one PR value per neuron.

Smoothing feature distributions

In Figures 4A, 5, 8-figure supplement 1 and 9G, distributions of features were smoothed with Kernel Density Estimation (KDE). Specifically, for a distribution, the density $\hat{f}_h$ at point *x* was computed as

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right)$$

where $\{x_1, x_2, x_3, \ldots, x_n\}$ represent the distribution of interest, *K* is the Gaussian kernel, and the bandwidth *h* was set by Silverman's rule (Läuter 1988):

$$h = 1.06\frac{A}{n^{1/5}}, A = \min\left(s, \frac{\mathrm{IQR}(x)}{1.34}\right)$$

with *n* being the number of observations *s* and being the sample standard deviation.

PCA-based alternative to feature-based analyses

As an alternative to our feature extraction method for summarizing population structure in trial-averaged activity, we applied PCA to the vectorized PETHs of all modulated neurons (the matrix of shape N x CT) and used the first 20 coefficients of each neuron as a feature space. For increased interpretability, we performed a VARIMAX rotation on the first 20 coefficients to sparsify them without distorting the space. The rotated scores' spatial structure was visualized in Figure 5-figure supplement 1. The likelihood maps for clustering based on the PC coefficients were shown in Figure 9-figure supplement 3C-G.

Gradient calculation

In Figure 6, we asked whether the single neuron PETH feature values changed more over the same spatial distance in some locations than others. This corresponds to a spatial derivative of the metric field. To compute this for each map in Figure 6, we first smoothed the spatial distribution of the feature map using the irregular Gaussian smooth function as described in the section "Computing map images from single-cell properties". We then gridded the smoothed map values into square bins as described above, and filled missing interior values with linear interpolations of their neighbors. At each pixel, we then computed the 2D gradient magnitude using MATLAB's centered finite-difference method. Specifically, if $M(x,y)$ denotes the gridded feature map, then the gradient magnitude $|\nabla M(x,y)|$ was calculated as:

$$|\nabla M(x,y)| = \sqrt{\left(\frac{\partial M(x,y)}{\partial x}\right)^2 + \left(\frac{\partial M(x,y)}{\partial y}\right)^2}$$

This produced a scalar map of local changes in the corresponding feature map. For the version in Figure 9-figure supplement 5, the two partial derivatives (x and y) were computed for each of the four subpopulation likelihood maps. The quadratic mean was then calculated as above over the set of individual feature (or likelihood) gradient magnitude maps.

Random boundary analysis

In Figure 6G,H we generated random boundaries through rejection sampling. Potential midpoint coordinates were randomly selected uniformly within the bounding box of the map, and an orientation was selected from a uniform distribution of 0-360°. We then computed the endpoints for a 1 mm line segment with that center and orientation. This boundary candidate was accepted if two criteria were met. First, both endpoints must have been within a non-empty pixel. Second, at least 200 neurons must have been in the valid zones on each side of the boundary (between 50 μm and 500 μm from the boundary). To account for non-uniform distributions of neurons, we downsampled neurons on the denser side of each random boundary to match the count on the sparser side.

To assess how well we could classify neurons as above or below a random boundary using the neurons' feature vectors, we trained a binary SVM with a radial basis function kernel and parameter optimization via Bayesian hyperparameter search. Decoding performance was evaluated using 5-fold cross-validation.

Nonlinear dimensionality reduction

We used t-SNE to visualize the structure of our neural population's feature space. Each neuron was represented by a five-dimensional feature vector comprising the PETH feature values described above. To improve the representation of large-scale structure, we seeded the t-SNE optimization with the top two principal components of the feature matrix (Kobak and Linderman 2021). The affinity metric was chosen to be the Pearson correlation. t-SNE perplexity hyperparameter was chosen according to the $N/100$ rule of thumb (Kobak and Berens 2019). The t-SNE space produced a two-dimensional projection of our neural data with coordinates of $Y_i = (y_{i,1}, y_{i,2})$ for each neuron i.

To produce contours in the t-SNE space for each area, we first selected the points for a given area and performed two-dimensional Gaussian KDE on these points to produce an estimate of probability density. We then drew contours (*contour* in Matlab) at the 90th, 95th, and 98th percentiles of the peak. To verify that our results were not specific to t-SNE, we also applied two-dimensional Uniform Manifold Approximation and Projection (UMAP; McInnes et al. (2020)) and obtained qualitatively similar results.

Gaussian Mixture Modeling and cluster selection

For the analyses in Figures 8, 9 and 10, we modeled the PETH feature vectors from each area using Gaussian Mixture Models (GMMs). For each candidate number of clusters *k*, we fit a GMM with full covariance matrices using the Expectation-Maximization algorithm (fitgmdist in Matlab). To select the optimal number of mixture components, we used the Integrated Completed Likelihood (ICL) criterion (Bertoletti et al. 2015), computed as:

$$\mathrm{ICL}(k) = -\mathrm{BIC}(k) - \sum_{i,j} P_{ij} \log\left(P_{ij} + \epsilon\right)$$

where P_{ij} denotes the posterior probability that neuron i belongs to cluster j, BIC represents the Bayesian Information Criterion (Schwarz 1978), and ϵ was chosen as 10^{-4}. For each area, we fitted GMMs with one to ten components, and selected the fit that produced the maximal ICL.

Likelihood computations

Following model selection for each region, we computed the log likelihood of every neuron according to the GMM of each anatomical region. The log-likelihood $\mathscr{L}_i$ of neuron i given the full model for a single area was defined as:

$$\mathscr{L}_i = \log\left(\sum_{j=1}^{K} \pi_j \mathcal{N}\left(X_i \mid \mu_j, \Sigma_j\right) + \epsilon\right)$$

where K is the number of components in the GMM and π_j is the mixing weight for component j. In addition to the full-mixture model, log-likelihoods were calculated for each component within an anatomical area's full mixture model to quantify the spatial footprint of each Gaussian component and to compare this spatial map across components of different anatomical areas. These log likelihoods provide a computational approach for understanding the structure of single neuron feature properties across the sensorimotor cortex.

GMM component distances and clustering

In Figure 9B, we computed the distance between each pair of GMM components found in different areas. To do so, we computed pairwise Bhattacharyya distances between the Gaussians. This distance captures differences in the means (as a Mahalanobis-like distance), and differences in the variances and covariances (differences in the orientations, shapes, and sizes). Each Gaussian k was represented by its mean vector $\boldsymbol{\mu}_k \in \mathbb{R}^N$ and covariance matrix $\sum_k \in \mathbb{R}^{N\times N}$, where N is the feature dimensionality (5). The Bhattacharyya distance D_B between multivariate Gaussians i and j was defined as

$$D_B\left(p_i, p_j\right) = \frac{1}{8}\left(\boldsymbol{\mu_i} - \boldsymbol{\mu_j}\right)^{\top} \Sigma^{-1} \left(\boldsymbol{\mu_i} - \boldsymbol{\mu_j}\right) + \frac{1}{2}\ln\left(\frac{\det(\Sigma)}{\sqrt{\det\left(\Sigma_j\right)\det\left(\Sigma_j\right)}}\right)$$

$$\text{where } \Sigma = \frac{1}{2}\left(\Sigma_i + \Sigma_j\right)$$

For display, the distance matrix produced by the Bhattacharyya distance was ordered using hierarchical clustering with Ward's method (Figure 9B). The number of clusters ('subpopulations') was determined by simple examination of this matrix. Each cluster comprised GMM components with heavily overlapping distributions in the feature space. Clusters were further validated by comparing the spatial distributions (in cortex) of the member neurons (Figure 10E), which were very similar for the components grouped in each cluster.

As the simplest way to fairly represent the components that were grouped to form subpopulations, for further analysis we modeled each subpopulation as arising from a GMM comprising the Gaussian components that were grouped together by clustering, weighted equally.

Data availability

Data will be made available on FigShare upon publication.

Acknowledgements

The authors thank P. Ravishankar for animal care assistance and L. Pinto for helpful comments on the manuscript. This work was funded by The University of Chicago, NIH-NINDS R01 NS121535 (MK), the Simons Foundation (MK), the Whitehall Foundation (MK), the NSF-Simons National Institute for Theory and Mathematics in Biology via grants NSF DMS-2235451 and Simons Foundation MP-TMPS-00005320 (MK), and NIH T32 NS121763 (HG).

Additional information

Code availability

Code used to create the figures is available on GitHub: *https://github.com/kaufmanlab/SGK25-public*.

Funding

Funder	Grant reference number	Author
National Science Foundation (NSF)	NCS 1835390	Matthew Tyler Kaufman
HHS | NIH | National Institute of Neurological Disorders and Stroke (NINDS)	R01 NS121535	Matthew Tyler Kaufman
Simons Foundation (SF)	876393SPI	Matthew Tyler Kaufman
Alfred P. Sloan Foundation (APSF)		Matthew Tyler Kaufman
Whitehall Foundation (The Whitehall Foundation)		Matthew Tyler Kaufman
National Science Foundation (NSF)	DMS-2235451	Matthew Tyler Kaufman
Simons Foundation (SF)	MP-TMPS-00005320	Matthew Tyler Kaufman
HHS | NIH | National Institute of Neurological Disorders and Stroke (NINDS)	T32 NS121763	Harrison Grier

Author ORCID iDs

Harrison A Grier: https://orcid.org/0000-0001-9107-2912
Matthew T Kaufman: https://orcid.org/0000-0002-8072-023X

Additional files

Supplemental figures Figure 1-figure supplement 1. Widefield window images and vibration maps aligned to the Allen CCF for all mice. Shown as in Figure 1B. Figure 2-figure supplement 1. Reach-to-grasp-to-drink centroid trajectories show moderate target-specificity. (A) Single-trial finger centroid trajectories locked to lift onset as in Figure 2C, shown from a posterior to anterior perspective, one session per mouse. Target positions not shown. Single-trials highlighted for one target for mouse 3, the same mouse shown in Figure 2C. (B) Each trace quantifies how

different reaches were on average. For each possible pair of trials, we computed the Euclidean distance between finger centroid XYZ positions at each time point. Red traces show the average distance across trials to different targets and black traces show the average distance across trials to the same target. Shading shows SEMs across pairwise distances. In many cases the shading is occluded by the mean trace. Horizontal bars above the traces show time points with significantly different distributions of within vs. between distances at $p < 0.05$ from a two sample Kolmogorov-Smirnov test. Figure 3-figure supplement 1. Modulation maps for all task events. (A) Binned and smoothed modulation map as in Figure 3B, but for cells modulated by Cue (left), Lift (center) or First-contact (right). Cue alignment used neural data from 200 ms before to 400 ms after cue. Lift alignment used neural data from 200 ms before to 400 ms after lift. First-contact alignment used neural data from 300 ms before to 300 ms after first spout contact. (B) Binary matrix indicating which area-area comparisons had significantly different proportions of cue-modulated cells. Black squares indicate rejection of the null hypothesis at $p < 0.005$ using a two proportion z-test. Figure 5-figure supplement 1. Principal component scores after VARIMAX rotation for each neuron. Plotted as feature maps using the same method as Figure 5. Figure 8-figure supplement 1. Feature value histograms for all GMM components and areas. Figure 9-figure supplement 1. Highest likelihood PETHs from each subpopulation. Two PETHs are shown from each subpopulation, for a total of four per area. The two PETHs from each subpopulation for each area had the highest likelihood values across cells from that area. PETHs were chosen in a principled way as the highest likelihood neurons for each subpopulation within each area, with a max firing rate of at least 33 events/s. This approach identified the only two Anterior cells in S1-hl and zero cells in S1-tr; thus the S1-tr entries were left blank. Figure 9-figure supplement 2: Separation of subpopulation distributions in feature space. (A) Feature vectors for all neurons within a subpopulation were projected onto the axis that best linearly separated the distributions, for each pairwise comparison between subpopulations. The axis was obtained from a logistic regression and the cross-validated performance of the model is shown for each comparison. Note that subpopulation membership was determined by GMM classification, which is nonlinear, while logistic regression is linear. Because of both this model mismatch and cross-validation of the logistic regression, classification performance on the projection axis is less than 100%. (B) t-SNE plot of feature vectors as in Figure 7 but colored by subpopulation membership. (C) Time to half-max distributions for each subpopulation, plotted as in Figure 4A. Figure 9-figure supplement 3. Additional subpopulation validation analyses. (A) t-SNE as in Figure 7C but using the top 20 PCs of the PETHs as inputs instead of feature vectors. Multimodality was again strongly present. (B) Left, subsampled population of neurons chosen from an overlap zone in M1, to analyze for discrete subpopulations while avoiding spatial somatotopy. *Left*, t-SNE embedding of this M1 population using the PETH feature space. *Right*, t-SNE embedding of the same M1 population using the PCs. Both methods yielded clear multimodality. (C-G) Same as Figure 8B-F but using top 20 PCs instead of PETH features. Figure 9-figure supplement 4. GMM fit to all PETH feature vectors together, agnostic to anatomical areas. One GMM was fit to the feature vectors of cells from all 5 main areas. Each map plots the likelihood for all cells to each of the three components of this area-agnostic GMM. Figure 9-figure supplement 5. Gradients of subpopulation likelihood maps. (A-D) Gradients as in Figure 6 but calculated using the subpopulation likelihood maps depicted in Figure 9C-F instead of the PETH feature maps. As in Figure 6, each subpopulation had a distinct gradient with boundaries approximately aligned with borders between anatomical regions, and other boundaries approximately separating somatotopic representations. (E) The quadratic mean of the four subpopulation gradient maps. This pooled gradient map closely resembles the pooled map in Figure 6F.

References

Allen ND (2025) aind-ophys-extraction. Github. https://github.com/AllenNeuralDynamics/aind-ophys-extraction

Alonso I, Scheer I, Palacio-Manzano M, Frézel-Jacob N, Philippides A, Prsa M. (2023) Peripersonal encoding of forelimb proprioception in the mouse somatosensory cortex. *Nat Commun* **14**:1866 https://doi.org/10.1038/s41467-023-37575-w | PubMed

An X, Li Y, Matho K, Mohan H, Xu XH, Whishaw IQ, Kepecs A, Huang ZJ (2025) A cortical circuit for orchestrating oromanual food manipulation. *bioRxiv* https://doi.org/10.1101/2022.12.03.518964

Bernardi S, Benna MK, Rigotti M, Munuera J, Fusi S, Salzman CD (2020) The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. *Cell* **183**:954-967.e21 https://doi.org/10.1016/j.cell.2020.09.031 | PubMed

Bertoletti M, Friel N, Rastelli R. (2015) Choosing the number of clusters in a finite mixture model using an exact Integrated Completed Likelihood criterion. *arXiv* https://doi.org/10.48550/arXiv.1411.4257

Brodmann K. (2005) *Brodmann's: Localisation in the Cerebral Cortex* (3rd) New York, NY: Springer.

Carmona LM, Thomas ED, Smith K, Tasic B, Costa RM, Nelson A. (2024) Topographical and cell type-specific connectivity of rostral and caudal forelimb corticospinal neuron populations. *Cell Rep* **43**:113993 https://doi.org/10.1016/j.celrep.2024.113993 | PubMed

Caro-Martín CR, Delgado-García JM, Gruart A, Sánchez-Campusano R. (2018) Spike sorting based on shape, phase, and distribution features, and K-TOPS clustering with validity and error indices. *Sci Rep* **8**:17796 https://doi.org/10.1038/s41598-018-35491-4 | PubMed

Chen JL, Carta S, Soldado-Magraner J, Schneider BL, Helmchen F. (2013) Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature* **499**:336-340 https://doi.org/10.1038/nature12236 | PubMed

Chen T-W, Li N, Daie K, Svoboda K. (2017) A Map of Anticipatory Activity in Mouse Motor Cortex. *Neuron* **94**:866-879.e4 https://doi.org/10.1016/j.neuron.2017.05.005 | PubMed

Costa RM, Luo JK, Salvino PS, Ackert-Smith LA, Ibarra SM, Pinto L. (2025) Cognitive processes are disentangled at cortex-wide scales. *bioRxiv* https://doi.org/10.1101/2025.07.24.666672 | PubMed

Currie SP, Ammer JJ, Premchand B, Dacre J, Wu Y, Eleftheriou C, Colligan M, Clarke T, Mitchell L, Faisal AA, *et al.* (2022) Movement-specific signaling is differentially distributed across motor cortex layer 5 projection neuron classes. *Cell Rep* **39**:110801 https://doi.org/10.1016/j.celrep.2022.110801 | PubMed

DeNardo LA, Berns DS, DeLoach K, Luo L. (2015) Connectivity of mouse somatosensory and prefrontal cortex examined with trans-synaptic tracing. *Nat Neurosci* **18**:1687-1697 https://doi.org/10.1038/nn.4131 | PubMed

Disse GD, Nandakumar B, Pauzin FP, Blumenthal GH, Kong Z, Ditterich J, Moxon KA (2023) Neural ensemble dynamics in trunk and hindlimb sensorimotor cortex encode for the control of postural stability. *Cell Rep* **42**:112347 https://doi.org/10.1016/j.celrep.2023.112347 | PubMed

Economo MN, Viswanathan S, Tasic B, Bas E, Winnubst J, Menon V, Graybuck LT, Nguyen TN, Smith KA, Yao Z, *et al.* (2018) Distinct descending motor cortex pathways and their roles in movement. *Nature* **563**:79-84 https://doi.org/10.1038/s41586-018-0642-9 | PubMed

Esmaeili V, Tamura K, Muscinelli SP, Modirshanechi A, Boscaglia M, Lee AB, Oryshchuk A, Foustoukos G, Liu Y, Crochet S, *et al.* (2021) Rapid suppression and sustained activation of distinct cortical regions for a delayed sensory-triggered motor response. *Neuron* **109**:2183-2201.e9 https://doi.org/10.1016/j.neuron.2021.05.005 | PubMed

Galiñanes GL, Bonardi C, Huber D. (2018) Directional Reaching for Water as a Cortex-Dependent Behavioral Framework for Mice. *Cell Rep* **22**:2767-2783 https://doi.org/10.1016/j.celrep.2018.02.042 | PubMed

Galiñanes GL, Huber D. (2023) Continuous estimation of reaching space in superficial layers of the motor cortex. *bioRxiv* https://doi.org/10.1101/2023.12.01.569527

Gămănuţ R, Kennedy H, Toroczkai Z, Ercsey-Ravasz M, Van Essen DC, Knoblauch K, Burkhalter A. (2018) The Mouse Cortical Connectome, Characterized by an Ultra-Dense Cortical Graph, Maintains Specificity by Distinct Connectivity Profiles. *Neuron* **97**:698-715.e10 https://doi.org/10.1016/j.neuron.2017.12.037 | PubMed

Grier HA, Salimian S, Kaufman MT (2026) Mouse sensorimotor cortex reflects complex kinematic details during reaching and grasping. *eLife* **14**:RP106270 https://doi.org/10.7554/eLife.106270 | PubMed

Guo J-Z, Graves AR, Guo WW, Zheng J, Lee A, Rodríguez-González J, Li N, Macklin JJ, Phillips JW, Mensh BD, *et al.* (2015) Cortex commands the performance of skilled movement. *eLife* **4**:e10774 https://doi.org/10.7554/eLife.10774 | PubMed

Harris JA, Hirokawa KE, Sorensen SA, Gu H, Mills M, Ng LL, Bohn P, Mortrud M, Ouellette B, Kidney J, *et al.* (2014) Anatomical characterization of Cre driver mice for neural circuit mapping and manipulation. *Front Neural Circuits* **8**:76 https://doi.org/10.3389/fncir.2014.00076 | PubMed

Harris JA, Mihalas S, Hirokawa KE, Whitesell JD, Choi H, Bernard A, Bohn P, Caldejon S, Casal L, Cho A, *et al.* (2019) Hierarchical organization of cortical and thalamic connectivity. *Nature* **575**:195-202 https://doi.org/10.1038/s41586-019-1716-z | PubMed

Hausmann FS, Barrett JM, Martin ME, Zhan H, Shepherd GMG (2022) Axonal Barcode Analysis of Pyramidal Tract Projections from Mouse Forelimb M1 and M2. *J Neurosci* **42**:7733-7743 https://doi.org/10.1523/jneurosci.1062-22.2022 | PubMed

Heindorf M, Arber S, Keller GB (2018) Mouse Motor Cortex Coordinates the Behavioral Response to Unpredicted Sensory Feedback. *Neuron* **99**:1040-1054.e5 https://doi.org/10.1016/j.neuron.2018.07.046 | PubMed

Hinton GE, Roweis S. (2022) Stochastic Neighbor Embedding. In: Advances in Neural Information Processing Systems.

Hira R, Ohkubo F, Ozawa K, Isomura Y, Kitamura K, Kano M, Kasai H, Matsuzaki M. (2013a) Spatiotemporal Dynamics of Functional Clusters of Neurons in the Mouse Motor Cortex during a Voluntary Movement. *J Neurosci* **33**:1377-1390 https://doi.org/10.1523/jneurosci.2550-12.2013 | PubMed

Hira R, Ohkubo F, Tanaka YR, Masamizu Y, Augustine GJ, Kasai H, Matsuzaki M. (2013b) In vivo optogenetic tracing of functional corticocortical connections between motor forelimb areas. *Front Neural Circuits* **7** https://doi.org/10.3389/fncir.2013.00055 | PubMed

Hubert L, Arabie P. (1985) Comparing partitions. *J Classif* **2**:193-218 https://doi.org/10.1007/bf01908075

Hwang EJ, Link TD, Hu YY, Lu S, EH-J Wang, Lilascharoen V, Aronson S, O'Neil K, Lim BK, Komiyama T. (2019) Corticostriatal Flow of Action Selection Bias. *Neuron* **104**:1126-1140.e6 https://doi.org/10.1016/j.neuron.2019.09.028 | PubMed

Inagaki HK, Inagaki M, Romani S, Svoboda K. (2018) Low-Dimensional and Monotonic Preparatory Activity in Mouse Anterior Lateral Motor Cortex. *J Neurosci* **38**:4163-4185 https://doi.org/10.1523/jneurosci.3152-17.2018 | PubMed

Jeong M, Kim Y, Kim J, Ferrante DD, Mitra PP, Osten P, Kim D. (2016) Comparative three-dimensional connectome map of motor cortical projections in the mouse brain. *Sci Rep* **6**:20072 https://doi.org/10.1038/srep20072 | PubMed

Kaufman MT, Benna MK, Rigotti M, Stefanini F, Fusi S, Churchland AK (2022) The implications of categorical and category-free mixed selectivity on representational geometries. *Curr Opin Neurobiol* **77**:102644 https://doi.org/10.1016/j.conb.2022.102644 | PubMed

Kaufman MT, Seely JS, Sussillo D, Ryu SI, Shenoy KV, Churchland MM (2016) The Largest Response Component in the Motor Cortex Reflects Movement Timing but Not Movement Type. *eNeuro* **3** https://doi.org/10.1523/eneuro.0085-16.2016 | PubMed

Kerr JND, de Kock CPJ, Greenberg DS, Bruno RM, Sakmann B, Helmchen F. (2007) Spatial Organization of Neuronal Population Responses in Layer 2/3 of Rat Barrel Cortex. *J Neurosci* **27**:13316-13328 https://doi.org/10.1523/jneurosci.2210-07.2007 | PubMed

Kiani R, Cueva CJ, Reppas JB, Peixoto D, Ryu SI, Newsome WT (2015) Natural Grouping of Neural Responses Reveals Spatially Segregated Clusters in Prearcuate Cortex. *Neuron* **85**:1359-1373 https://doi.org/10.1016/j.neuron.2015.02.014 | PubMed

Kim M-H, Znamenskiy P, Iacaruso MF, Mrsic-Flogel TD (2018) Segregated Subnetworks of Intracortical Projection Neurons in Primary Visual Cortex. *Neuron* **100**:1313-1321.e6 https://doi.org/10.1016/j.neuron.2018.10.023 | PubMed

Kobak D, Berens P. (2019) The art of using t-SNE for single-cell transcriptomics. *Nat Commun* **10**:5416 https://doi.org/10.1038/s41467-019-13056-x | PubMed

Kobak D, Linderman GC (2021) Initialization is critical for preserving global data structure in both t-SNE and UMAP. *Nat Biotechnol* **39**:156-157 https://doi.org/10.1038/s41587-020-00809-z | PubMed

Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A. (2016) Correlations and Neuronal Population Information. *Annu Rev Neurosci* **39**:237-256 https://doi.org/10.1146/annurev-neuro-070815-013851 | PubMed

Komiyama T, Sato TR, O'Connor DH, Zhang Y-X, Huber D, Hooks BM, Gabitto M, Svoboda K. (2010) Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature* **464**:1182-1186 https://doi.org/10.1038/nature08897 | PubMed

Kovesi P. (2015) Good Colour Maps: How to Design Them. *arXiv* https://doi.org/10.48550/arxiv.1509.03700

Kriegeskorte N, Mur M, Bandettini PA (2008) Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci* **2** https://doi.org/10.3389/neuro.06.004.2008 | PubMed

Kristl AC, Koh N, Agrios M, Savya S, Ma Z, Basrai D, Hsu S, Miri A. (2025) Interactions between motor cortical forelimb regions and their influence on muscles reorganize across behaviors. *Nat Commun* **16**:10364 https://doi.org/10.1038/s41467-025-65381-z | PubMed

Läuter H. (1988) Silverman, B. W.: Density Estimation for Statistics and Data Analysis. *Biom J* **30**:876-877 https://doi.org/10.1002/bimj.4710300745

Lee EK, Balasubramanian H, Tsolias A, Anakwe SU, Medalla M, Shenoy KV, Chandrasekaran C. (2021) Non-linear dimensionality reduction on extracellular waveforms reveals cell type diversity in premotor cortex. *eLife* **10**:e67490 https://doi.org/10.7554/eLife.67490 | PubMed

Lewicki MS (1994) Bayesian Modeling and Classification of Neural Signals. *Neural Comput* **6**:1005-1030 https://doi.org/10.1162/neco.1994.6.5.1005

Li N, Chen T-W, Guo ZV, Gerfen CR, Svoboda K. (2015) A motor cortex circuit for motor planning and movement. *Nature* **519**:51-56 https://doi.org/10.1038/nature14178 | PubMed

Li Y, An X, Mulcahey PJ, Qian Y, Xu XH, Zhao S, Mohan H, Suryanarayana SM, Bachschmid-Romano L, Brunel N, *et al.* (2024) Cortico-thalamic communication for action coordination in a skilled motor sequence. *bioRxiv* https://doi.org/10.1101/2023.10.25.563871 | PubMed

Majka P, Bednarek S, Chan JM, Jermakow N, Liu C, Saworska G, Worthy KH, Silva AC, Wójcik DK, Rosa MGP (2021) Histology-Based Average Template of the Marmoset Cortex With Probabilistic Localization of Cytoarchitectural Areas. *NeuroImage* **226**:117625 https://doi.org/10.1016/j.neuroimage.2020.117625 | PubMed

Masamizu Y, Tanaka YR, Tanaka YH, Hira R, Ohkubo F, Kitamura K, Isomura Y, Okada T, Matsuzaki M. (2014) Two distinct layer-specific dynamics of cortical ensembles during learning of a motor task. *Nat Neurosci* **17**:987-994 https://doi.org/10.1038/nn.3739 | PubMed

Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M. (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21**:1281-1289 https://doi.org/10.1038/s41593-018-0209-y | PubMed

Mathis MW, Mathis A, Uchida N. (2017) Somatosensory Cortex Plays an Essential Role in Forelimb Motor Adaptation in Mice. *Neuron* **93**:1493-1503.e6 https://doi.org/10.1016/j.neuron.2017.02.049 | PubMed

Maurer L, Brown M, Saggi T, Cardiges A, Kolarcik CL (2023) Hindlimb muscle representations in mouse motor cortex defined by viral tracing. *Front Neuroanat* **17** https://doi.org/10.3389/fnana.2023.965318 | PubMed

Mayrhofer JM, El-Boustani S, Foustoukos G, Auffret M, Tamura K, Petersen CCH (2019) Distinct Contributions of Whisker Sensory Cortex and Tongue-Jaw Motor Cortex in a Goal-Directed Sensorimotor Transformation. *Neuron* **103**:1034-1043.e5 https://doi.org/10.1016/j.neuron.2019.07.008 | PubMed

McInnes L, Healy J, Melville J. (2020) UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* https://doi.org/10.48550/arXiv.1802.03426

Minderer M, Brown KD, Harvey CD (2019) The Spatial Structure of Neural Encoding in Mouse Posterior Cortex during Navigation. *Neuron* **102**:232-248.e11 https://doi.org/10.1016/j.neuron.2019.01.029 | PubMed

Mitz AR, Wise SP (1987) The somatotopic organization of the supplementary motor area: intracortical microstimulation mapping. *J Neurosci* **7**:1010-1021 https://doi.org/10.1523/jneurosci.07-04-01010.1987 | PubMed

Mohan H, An X, Xu XH, Kondo H, Zhao S, Matho KS, Wang B-S, Musall S, Mitra P, Huang ZJ (2023) Cortical glutamatergic projection neuron types contribute to distinct functional subnetworks. *Nat Neurosci* **26**:481-494 https://doi.org/10.1038/s41593-022-01244-w | PubMed

Montijn JS, Seignette K, Howlett MH, Cazemier JL, Kamermans M, Levelt CN, Heimel JA (2021) A parameter-free statistical test for neuronal responsiveness. *eLife* **10**:e71969 https://doi.org/10.7554/eLife.71969 | PubMed

Morandell K, Huber D. (2017) The role of forelimb motor cortex areas in goal directed action in mice. *Sci Rep* **7**:15759 https://doi.org/10.1038/s41598-017-15835-2 | PubMed

Muñoz-Castañeda R, Zingg B, Matho KS, Chen X, Wang Q, Foster NN, Li A, Narasimhan A, Hirokawa KE, Huo B, *et al.* (2021) Cellular anatomy of the mouse primary motor cortex. *Nature* **598**:159-166 https://doi.org/10.1038/s41586-021-03970-w | PubMed

Musall S, Kaufman MT, Juavinett AL, Gluf S, Churchland AK (2019) Single-trial neural dynamics are dominated by richly varied movements. *Nat Neurosci* **22**:1677-1686 https://doi.org/10.1038/s41593-019-0502-4 | PubMed

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

Nandakumar B, Blumenthal GH, Pauzin FP, Moxon KA (2021) Hindlimb Somatosensory Information Influences Trunk Sensory and Motor Cortices to Support Trunk Stabilization. *Cereb Cortex* **31**:5165-5187 https://doi.org/10.1093/cercor/bhab150 | PubMed

Neafsey EJ, Bold EL, Sievert CF, Terreberry RR, Quirk GJ (1983) The primary motor cortex of the owl monkey: a microstimulation mapping study of the leg and arm areas. *Soc Neurosci Abstr*

Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, Wang Q, Lau C, Kuan L, Henry AM, *et al.* (2014) A mesoscale connectome of the mouse brain. *Nature* **508**:207-214 https://doi.org/10.1038/nature13186 | PubMed

Park J, Phillips JW, Guo J-Z, Martin KA, Hantman AW, Dudman JT (2022) Motor cortical output for skilled forelimb movement is selectively distributed across projection neuron classes. *Sci Adv* **8**:eabj5167 https://doi.org/10.1126/sciadv.abj5167 | PubMed

Posani L, Wang S, Muscinelli SP, Paninski L, Fusi S. (2025) Rarely categorical, always high-dimensional: how the neural code changes along the cortical hierarchy. *bioRxiv* https://doi.org/10.12751/nncn.bc2025.012

Qi H-X, Jain N, Collins CE, Lyon DC, Kaas JH (2010) Functional organization of motor cortex of adult macaque monkeys is altered by sensory loss in infancy. *Proc Natl Acad Sci* **107**:3192-3197 https://doi.org/10.1073/pnas.0914962107 | PubMed

Saiki-Ishikawa A, Agrios M, Savya S, Forrest A, Sroussi H, Hsu S, Basrai D, Xu F, Miri A. (2025) Hierarchy between forelimb premotor and primary motor cortices and its manifestation in their firing patterns. *eLife* **13**:RP103069 https://doi.org/10.7554/eLife.103069 | PubMed

Saxena S, Kinsella I, Musall S, Kim SH, Meszaros J, Thibodeaux DN, Kim C, Cunningham J, Hillman EMC, Churchland A, *et al.* (2020) Localized semi-nonnegative matrix factorization (LocaNMF) of widefield calcium imaging data. *PLOS Comput Biol* **16**:e1007791 https://doi.org/10.1371/journal.pcbi.1007791 | PubMed

Schwarz G. (1978) Estimating the Dimension of a Model. *Ann Stat* **6**:461-464 https://doi.org/10.1214/aos/1176344136

Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD (2019) Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**:266-273 https://doi.org/10.1038/s41586-019-1787-x | PubMed

Stringer C, Pachitariu M. (2019) Computational processing of neural recordings from calcium imaging data. *Curr Opin Neurobiol* **55**:22-31 https://doi.org/10.1016/j.conb.2018.11.005 | PubMed

Sweeney CG, Thomas ME, Liu CJ, Vattino LG, Smith KE, Takesian AE (2025) Reliable sensory processing of superficial cortical interneurons is modulated by behavioral state. *Cell Rep* **44**:115678 https://doi.org/10.1016/j.celrep.2025.115678 | PubMed

Tennant KA, Adkins DL, Donlan NA, Asay AL, Thomas N, Kleim JA, Jones TA (2011) The Organization of the Forelimb Representation of the C57BL/6 Mouse Motor Cortex as Defined by Intracortical Microstimulation and Cytoarchitecture. *Cereb Cortex* **21**:865-876 https://doi.org/10.1093/cercor/bhq159 | PubMed

Terada S-I, Kobayashi K, Matsuzaki M. (2022) Transition of distinct context-dependent ensembles from secondary to primary motor cortex in skilled motor performance. *Cell Rep* **41** https://doi.org/10.1016/j.celrep.2022.111494 | PubMed

Ueno M, Nakamura Y, Li J, Gu Z, Niehaus J, Maezawa M, Crone SA, Goulding M, Baccei ML, Yoshida Y. (2018) Corticospinal Circuits from the Sensory and Motor Cortices Differentially Regulate Skilled Movements through Distinct Spinal Interneurons. *Cell Rep* **23**:1286-1300.e7 https://doi.org/10.1016/j.celrep.2018.03.137 | PubMed

Van Acker GM, Amundsen SL, Messamore WG, Zhang HY, Luchies CW, Kovac A, Cheney PD (2013) Effective intracortical microstimulation parameters applied to primary motor cortex for evoking forelimb movements to stable spatial end points. *J Neurophysiol* **110**:1180-1189 https://doi.org/10.1152/jn.00172.2012 | PubMed

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

Wang T, Pachitariu M, Guo C, Lee AK, Lee JS, Zhang Y, Rozsa M, Hasseman J, Kolb I, Cox JD, *et al.* (2023a) jGCaMP8 transgenic mice. Janelia Research Campus. https://doi.org/10.25378/janelia.24565837.v3

Wang Z, Maunze B, Wang Y, Tsoulfas P, Blackmore MG (2018) Global Connectivity and Function of Descending Spinal Input Revealed by 3D Microscopy and Retrograde Transduction. *J Neurosci* **38**:10566-10581 https://doi.org/10.1523/jneurosci.1196-18.2018 | PubMed

Wang ZA, Chen S, Liu Y, Liu D, Svoboda K, Li N, Druckmann S. (2023b) Not everything, not everywhere, not all at once: a study of brain-wide encoding of movement. *bioRxiv* https://doi.org/10.1101/2023.06.08.544257 | PubMed

Winnubst J, Bas E, Ferreira TA, Wu Z, Economo MN, Edson P, Arthur BJ, Bruns C, Rokicki K, Schauder D, *et al.* (2019) Reconstruction of 1,000 Projection Neurons Reveals New Cell Types and Organization of Long-Range Connectivity in the Mouse Brain. *Cell* **179**:268-281.e13 https://doi.org/10.1016/j.cell.2019.07.042 | PubMed

Yang W, Kanodia H, Arber S. (2023) Structural and functional map for forelimb movement phases between cortex and medulla. *Cell* **186**:162-177.e18 https://doi.org/10.1016/j.cell.2022.12.009 | PubMed

Yeterian EH, Pandya DN, Tomaiuolo F, Petrides M. (2012) The Cortical Connectivity of the Prefrontal Cortex in the Monkey Brain. *Cortex J Devoted Study Nerv Syst Behav* **48**:58-81 https://doi.org/10.1016/j.cortex.2011.03.004 | PubMed

Zatka-Haas P, Steinmetz NA, Carandini M, Harris KD (2021) Sensory coding and the causal impact of mouse cortex in a visual decision. *eLife* **10**:e63163 https://doi.org/10.7554/eLife.63163 | PubMed

Zhang Y, Rózsa M, Liang Y, Bushey D, Wei Z, Zheng J, Reep D, Broussard GJ, Tsang A, Tsegaye G, *et al.* (2023) Fast and sensitive GCaMP calcium indicators for imaging neural populations. *Nature* **615**:884-891 https://doi.org/10.1038/s41586-023-05828-9 | PubMed

Zhu F, Grier HA, Tandon R, Cai C, Agarwal A, Giovannucci A, Kaufman MT, Pandarinath C. (2022) A deep learning framework for inference of single-trial neural population dynamics from calcium imaging with subframe temporal resolution. *Nat Neurosci* **25**:1724-1734 https://doi.org/10.1038/s41593-022-01189-0 | PubMed

Zingg B, Hintiryan H, Gou L, Song MY, Bay M, Bienkowski MS, Foster NN, Yamashita S, Bowman I, Toga AW, *et al.* (2014) Neural Networks of the Mouse Neocortex. *Cell* **156**:1096-1111 https://doi.org/10.1016/j.cell.2014.02.023 | PubMed

Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**:140-143 https://doi.org/10.1038/370140a0 | PubMed

Peer reviews

Reviewer #1 (Public review):

Summary:

Here the authors address the organization of reach-related activity in layer 2/3 across a broad swath of anterodorsal neocortex that included large subregions of M1, M2, and S1. In mice performing a novel variant water-reaching task, the authors measured activity using two-photon fluorescence imaging of a GECI expressed in excitatory projection neurons. The authors found a substantial diversity of response patterns using a number of metrics they developed for characterizing the PETHs of neurons across reach conditions (target locations). By mapping single-neuron properties across cortex, the authors found substantial spatial variation, only some of which aligned with traditional boundaries between cortical regions. Using Gaussian mixture models, the authors found evidence of distinct response types in each region, with several types prominent in multiple cortical regions. Aggregating across regions, four primary subpopulations were apparent, each distinct in their average response properties. Strikingly, each subpopulation was observed in multiple regions, but subpopulation members from different regions exhibited largely similar response properties.

Strengths:

The work addresses a fundamental question in the field that has not previously been addressed at cellular resolution across such a broad cortical extent. I see this as truly foundational work that will support future investigation of how the rodent brain drives and controls reaching.

The quantification is thoughtful and rigorous. It is great that the authors provide explanation for and intuition behind their response metrics, rather than burying everything in the Methods.

The Discussion and general contextualization of the Results is thorough, thoughtful, and strong. It is great that the authors avoid the common over-interpretation of classical

observations regarding cortical organization that are endemic in the field.

All things considered, this is the best paper regarding spatial structure in the motor system I have ever read. The breadth of cellular resolution activity measurement, the rigor of the quantification, and the clear and open-minded interrogation of the data collectively have produced a very special piece of work.

Weaknesses:

There are two important issues left unaddressed that the authors plan to address in their future work. The first is the relation between observed neural activity patterns and movement kinematics, and in particular how much the activity variation across target locations may relate to the kinematic differences across these different conditions, as opposed to true higher-order movement features like reach direction. The second issue is how to interpret the results in relation to existing ideas about behavioral organization in motor/premotor cortex.

Comments on revised version:

The authors have done an excellent job addressing my previous concerns. I have no additional concerns with the manuscript.

https://doi.org/10.7554/eLife.109240.2.sa2

Reviewer #2 (Public review):

Summary:

The functional parcellation of cortical areas is a critical question in neuroscience. This is particularly true in frontal areas in mice. While sensory areas are relatively well characterized by their tuning to sensory stimuli, the situation is much less clear for motor areas. This has become even more ambiguous since recent studies using large-scale neuronal recordings consistently report mixed sensory and motor-related activity throughout the brain and motor mapping studies have shown that movements evoked by cortical stimulation are by no means limited to motor areas alone. Here, the authors use a correlation approach combining large-scale functional imaging at cellular-resolution with movement-tracking in mice executing a reaching task. Across multiple recording sessions in the same animals, the authors have imaged a large portion of the sensorimotor cortex at cellular resolution in mice performing a reaching task, recording the activity of nearly 40,000 neurons. By aligning the calcium signal of each neuron to three task events-the Go cue triggering the reach, the onset of paw lift, and the contact between the paw and the target-for different target positions, the authors identified different response patterns distributed differently across cortical areas. They defined a set of features that describe the neurons' response pattern, representing the temporal dynamics and tuning properties for the different target positions. These features were used to construct cortical maps, and the authors show that, interestingly, gradient maps obtained from the first derivative of the feature maps reveal sharp discontinuities at the boundaries between anatomically defined cortical areas. Using dimensionality reduction of the neuronal response features, the authors found that, despite clear differences in their average response properties, individual neurons from the same cortical areas do not form distinct clusters in the reduced-dimensional space. In fact, most areas contain heterogeneous neuronal populations, and most neuronal populations are present in multiple areas, albeit in different proportions. Interestingly, the authors identified four neuronal subpopulations based on the distance between the components of the Gaussian mixture model used to model the distribution of neurons within each area. One of these subpopulations is almost exclusively represented in the anterior M2 cortex, while another is broadly distributed across the different areas.

Strengths:

This article is based on an impressive dataset of nearly 40,000 neurons covering a large portion of the sensorimotor cortex and on innovative analytical approaches. This study is likely the first to clearly demonstrate boundaries between cortical areas defined based on the responses of individual neurons. This innovative approach to functional mapping of cortical areas potentially opens up new perspectives for higher-resolution mapping of frontal cortical areas, using a broader repertoire of sensory and motor evoked responses.

Weaknesses:

One limitation of this study - inherent in most cell imaging studies - is that it only takes into account the activity of neurons in superficial cortical layers. One might think that taking into account neuronal activity across the different layers would allow for an even finer functional cortical segmentation.

Comments on revised version:

The authors have answered all my questions and this new version has largely improved in clarity.

https://doi.org/10.7554/eLife.109240.2.sa1

Author response:

The following is the authors' response to the original reviews.

In preparation for release of the analysis code used in the paper, we made many analyses more parallel to one another in their exact preprocessing. This resulted in very slight changes to many panels, but these changes are nearly invisible and conclusions did not change. In one case, though, we realized that the way we were presenting data was potentially misleading (the timing plot in Figure 3A). The original plot was of the distribution of pixel values from the spatially smoothed map instead of distributions over individual neurons. We have now swapped it out for better interpretability and changed the accompanying text accordingly.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *Here, the authors address the organization of reach-related activity in layer 2/3 across a broad swath of anterodorsal neocortex that included large subregions of M1, M2, and S1. In mice performing a novel variant water-reaching task, the authors measured activity using two-photon fluorescence imaging of a GECI expressed in excitatory projection neurons. The authors found a substantial diversity of response patterns using a number of metrics they developed for characterizing the PETHs of neurons across reach conditions (target locations). By mapping single-neuron properties across the cortex, the authors found substantial spatial variation, only some of which aligned with traditional boundaries between cortical regions. Using Gaussian mixture models, the authors found evidence of distinct response types in each region, with several types prominent in multiple cortical regions. Aggregating across regions, four primary subpopulations were apparent, each distinct in its average response properties. Strikingly, each subpopulation was observed in multiple regions, but subpopulation members from different regions exhibited largely similar response properties.*
>
> *Strengths:*

> *The work addresses a fundamental question in the field that has not previously been addressed at cellular resolution across such a broad cortical extent. I see this as truly foundational work that will support future investigation of how the rodent brain drives and controls reaching.*
>
> *The quantification is thoughtful and rigorous. It is great that the authors provide an explanation for and intuition behind their response metrics, rather than burying everything in the Methods.*
>
> *The Discussion and general contextualization of the results are thorough, thoughtful, and strong. It is great that the authors avoid the common over-interpretation of classical observations regarding cortical organization that are endemic in the field.*
>
> *All things considered, this is the best paper regarding spatial structure in the motor system I have ever read. The breadth of cellular resolution activity measurement, the rigor of the quantification, and the clear and open-minded interrogation of the data collectively have produced a very special piece of work.*

Thank you! We really, really appreciate this!

> *Weaknesses:*
>
> *The behavioral task is very impressive and an important contribution to the field in its own right. However, given that it appears substantially different from the one used in the previous paper, the characterization of the behavior provided in the Results is too brief. More illustration of the behavior would be helpful. For example, it is rather deep into the paper when the authors reveal that the mice can whisk to help localize the target location. That should be expressed at the outset when the behavior is first described. Other suggestions for elaborating the behavior description are included below.*

Thank you. Although the task will be treated in greater detail in the next paper (where we more closely relate neural activity to the kinematics), we have added more exposition of the task here. In particular, we now include a figure with a characterization of the trial-to-trial variability across reaches to the same target versus across reaches to different targets (Figure 2-figure supplement 1B). This supports the idea that the mice aimed their reaches. We have also expanded that text.

Regarding whisking, we have now revised that text to make clear that we do not know how the mice localize the spout. The original work by Galinanes and Huber argued that they find the spout by sniffing the water; they may do the same here, or may find it via whisking. It is also possible that the whisking they do is simply because the spout moves in and they are excited, or startled, or do it by reflex. We simply have no evidence one way or another. We have therefore revised the text to make it clearer that whisking-related activation could have occurred for a variety of reasons.

> *Statistical support for key claims is lacking. For example, "The five areas of interest varied in the fraction of neurons that were modulated: M2 had 14%, M1 had 23%, S1-fl had 30%, S1-hl had 25%, and S1-tr had 27%" - I cannot locate the statistical tests showing that these values are actually different. Another example is Figure 7, where a key observation is that distributions of PETH features are distinct across regions. It is clear that at least some distributions are not overlapping, but a clearer statistical basis for this key claim should be provided.*

Good idea. For the proportions, we have now added first a Chi-square test for homogeneity to show that there is variation in the proportions, then shown the results of pairwise two-proportion Z tests (Bonferroni-corrected for multiple comparisons) as a binary matrix in

Figure 3-figure supplement 1B. For the area distributions in the t-SNE space (Figure 7), we have added a 2-dimensional Kolmogorov-Smirnov test, again corrected for multiple comparisons, with p-values quoted in the text.

> *I understand that the authors are planning a follow-up study that addresses the relation between activity patterns and kinematics. One question about interpreting the results here though, is how much the activity variation across target locations may relate to the kinematic differences across these different conditions, as opposed to true higher-order movement features like reach direction.*

We agree this is a very important question. However, having done many of the analyses to examine the question for the next paper in the series, we do not know of a shortcut to the right answer. This question requires thorough treatment, and so we leave it to be covered in subsequent work. Instead, after our speculation about how responses suggest function, we are now explicit that these hypotheses needs testing:

"In each of these cases, determining the relationships of the observed activity patterns to function will require specific attempts to link the activity to kinematics, target location, sensory feedback, and more; these relationships will be addressed in future work."

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *The functional parcellation of cortical areas is a critical question in neuroscience. This is particularly true in frontal areas in mice. While sensory areas are relatively well characterized by their tuning to sensory stimuli, the situation is much less clear for motor areas. This has become even more ambiguous since recent studies using large-scale neuronal recordings consistently report mixed sensory and motor-related activity throughout the brain, and motor mapping studies have shown that movements evoked by cortical stimulation are by no means limited to motor areas alone. Here, the authors use a correlation approach combining large-scale functional imaging at cellular resolution with movement-tracking in mice executing a reaching task. Across multiple recording sessions in the same animals, the authors have imaged a large portion of the sensorimotor cortex at cellular resolution in mice performing a reaching task, recording the activity of nearly 40,000 neurons. By aligning the calcium signal of each neuron to three task events-the Go cue triggering the reach, the onset of paw lift, and the contact between the paw and the target-for different target positions, the authors identified different response patterns distributed differently across cortical areas. They defined a set of features that describe the neurons' response pattern, representing the temporal dynamics and tuning properties for the different target positions. These features were used to construct cortical maps, and the authors show that, interestingly, gradient maps obtained from the first derivative of the feature maps reveal sharp discontinuities at the boundaries between anatomically defined cortical areas. Using dimensionality reduction of the neuronal response features, the authors found that, despite clear differences in their average response properties, individual neurons from the same cortical areas do not form distinct clusters in the reduced-dimensional space. In fact, most areas contain heterogeneous neuronal populations, and most neuronal populations are present in multiple areas, albeit in different proportions. Interestingly, the authors identified four neuronal subpopulations based on the distance between the components of the Gaussian mixture model used to model the distribution of neurons within each area. One of these subpopulations is almost exclusively represented in the anterior M2 cortex, while another is broadly distributed across the different areas.*
>
> *Strengths:*

> *This article is based on an impressive dataset of nearly 40,000 neurons covering a large portion of the sensorimotor cortex and on innovative analytical approaches. This study is likely the first to clearly demonstrate boundaries between cortical areas defined based on the responses of individual neurons. This innovative approach to functional mapping of cortical areas potentially opens up new perspectives for higher-resolution mapping of frontal cortical areas, using a broader repertoire of sensory and motor evoked responses.*

Thank you!

> *Weaknesses:*
>
> *The second part of the article, which presents multimodal responses in the cortical areas, seems to be a perhaps overly complicated way of showing what has already been demonstrated in numerous recent publications, but these new analyses expand upon these previous observations by revealing an interesting functional organization of the sensorimotor cortex, highlighting interesting similarities and differences between certain areas.*

We understand the concern: a number of recent papers have also noted different neuron response characteristics distributed throughout the motor system. We compare and contrast in greater detail following the more specific comments on this below, but we briefly summarize here. The way previous work handled the data – for example, starting with PCA – mixes what neurons are tuned for and when they are tuned for it with what we refer to as the "response format": properties like tuning sharpness, response duration, etc. We focused primarily on this response format, and designed our features to be mostly independent of tuning preferences or peak response timing. We therefore pick up on different properties of neurons' responses than those prior works. In addition, no previous work we know of examined these properties across large swathes of cortex at single-cell resolution in the context of forelimb control. Together, these aspects of our work allowed us to produce high-resolution mapping of response properties in a way we have not seen in any prior work.

> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *In addition to addressing the weaknesses stated above, I suggest the authors also consider the following.*
>
> *The one big question left unresolved here is whether we should be thinking about these four subpopulations as distinct types with a biological basis and importance, or just reflections of activity pattern heterogeneity. The authors say that "we did not observe tight clusters in feature space separated by gaps," but their discussion here is light and a bit unclear, and their engagement with the issue of types versus heterogeneity, in my view, could be improved. We do not need "gaps" where the density goes to zero in parameter space, but we do need reproducible troughs between peaks. The authors should clarify if there are substantial and reproducible troughs in the parameter space between their four subpopulations.*

This is a great idea, and we have added three analyses and additional text to address it. We break this concern down into two more specific questions, based on the next comment by this reviewer.

(1) Are the clusters well separated / do they have troughs between them? (Note that even with troughs, clustering might not be stable if the clustering algorithm is poorly matched to the shapes of the clusters.)

(2) Is the clustering stable? (It can be stable even without troughs, if, for example, the distribution has a long tail and a GMM needs one Gaussian for the body of the distribution and a second for the tail.)

First, to directly address the presence or absence of troughs between clusters, we have added Figure 9-figure supplement 2A and 2B. For each pair of subpopulations, we trained a logistic regression classifier to separate the 5D feature vectors of the neurons in one subpopulation from the feature vectors of the neurons in the other subpopulation, then projected the feature vectors onto this axis. Note that because the subpopulations are defined by GMMs, which have nonlinear boundaries, the (linear) logistic classifier does not typically produce perfect classification. Nevertheless, this analysis provides a window onto how well separated each cluster is from each other cluster in feature space. In 5 of the 6 pairwise comparisons, it is obvious that the distributions are different and have at least some dip in the distribution density at the boundary. The one pair of clusters without a trough between them were the forelimb somatomotor and hindlimb somatomotor subpopulations. This was surprising to us, given that their likelihood maps are so strongly distinct, but this presumably reflects trying to capture a nonlinear classifier boundary with a linear one (see below). Overall, this analysis argues that the clusters do have fuzzy edges that blend into one another, but reflect concentration of mass near the centers of the clusters we identified.

Second, to address the same question with a different nonlinear method, we have added a version of the t-SNE plot from Figure 7 that is instead colored and contoured by subpopulation identity instead of area (Figure 9-figure supplement 2B). Agreement with the GMMs is not a given here either, because t-SNE is a fundamentally different and independent nonlinear transform from that performed by the GMM classification. Nevertheless, the subpopulations were again nicely separated – though not with troughs, possibly thanks to the inherent difficulty of interpreting point density with t-SNE. Interestingly, here the hindlimb somatomotor subpopulation was the best separated from the other subpopulations, supporting the idea that the lack of separation we observed above with the logistic projections was indeed due to a nonlinear boundary. This analysis again argues that neurons are more likely to have features that lie near the center of a cluster, but that the edges of the clusters run into one another. Additionally, this analysis makes clear that treating the hindlimb somatomotor subpopulation as a second cluster can be supported by other analyses, even if not by the logistic regression projection.

Third, to address the question of cluster stability, we have performed random splits of our data, GMM clustered the two halves independently, applied the GMM from one half to the other, and asked how similar the clusterings are using the Adjusted Rand Index. This produced a value of 0.856, which for this sensitive measure argues that the clustering is rather stable (at least for the three clusters that can be found with all data together, which does not include the smaller-in-size Anterior subpopulation). Note that we did not perform this analysis on the more complicated version where we fit a GMM to each area separately then cluster those; in our main analysis, the hierarchical clustering agreed with what we found by eye, but determining the number of clusters for hierarchical clustering is in general very unstable and so we did not have an objective way to determine the “true” number of clusters.

In addition to these new analyses, we note that three analyses we had already included bore strongly on this issue. Regarding separability of the clusters, the fact that our likelihood maps (Figure 9C-F) were quite distinct for different subpopulations argues that we picked up on ‘real’ differences. Second, Figure 9B found that when clustering non-overlapping data – different cells from different areas – we obtained clusters that were nearly identical in their feature distributions. Third, Figure 10E used the clusterings from different areas’ data to create likelihood maps, and found that they were extremely similar. These analyses together argue strongly that we are finding ‘types’ in a meaningful sense; given that we know the

areas do have different distributions of properties, if there weren't types then clustering would yield different clusters for different areas. Given the importance of the question, however, we are grateful that the reviewer encouraged us to find additional ways to make this point!

> *The original t-SNE plot is beautiful and quasi-fractalic, but it does not show clear signs of four cell types. The single-neuron activity profiles are clearly heterogeneous in very interesting ways, but heterogeneous does not imply a strong or reproducible multimodality that would indicate meaningful cell types. Clustering algorithms will always spit out an answer. If you just have elements uniformly distributed across a parameter space, plus some noise, when you ask for X clusters, you will get X clusters that have different centroids. When you ask an algorithm to cluster without defining the number of clusters, noise can lead the algorithm to produce a particular number of clusters that again will have distinct centroids. The salient question, though, is whether in the present case there is a parameter space in which the clusters are substantially and or reproducibly distinct. Distinct here would mean that peaks in the density across some parameter space are separated by troughs - again, we don't need true gaps. The more substantial the differences between clusters are (again, not the differences between centroids but the prominence of the density troughs between them), the more biologically meaningful the clustering is likely to be. Reproducibility here could be addressed with resampling methods (e.g., how often do two separate halves of the cells produce the same clusters?).*

Please see the reply above, which includes our addressing of this concern.

> *The Introduction is generally good, but it could further develop existing ideas about how function is distributed across cell types and regions. We would like to be able to imagine different answers to the question of how activity patterns are organized that might have divergent implications for how the circuit works. I understand we have very little to go on in terms of data, but I think it would be helpful for readers to be given more of a sense of what *could* be important.*

Good idea. We have added such a paragraph to the Introduction:

"To frame possible outcomes, consider that single neuron responses can vary along many dimensions. Cells could differ according to which movements or time periods they are recruited for (tuning), what movement parameters their activities reflect (encoding), or how their responses are structured across different movements (e.g., nonlinear encoding structure). Further, differences in these response properties across cells could be distributed over the cortical sheet in a variety of ways. Cells could form distinct "categories" or clusters that are spatially well-aligned to the boundaries of anatomically defined regions. Or, categories of neurons might span area boundaries in spatial footprints that do not relate obviously to area boundaries, and that either abut or overlap. At a fine-grained scale, cells with similar responses could be physically located near one another as in primate and feline visual cortex, or similarly-responsive neurons might be salt-and-pepper intermingled as seen in rodent visual cortex or in primate motor cortices during reaching behaviors."

> *It should be clarified in the Results how the cue relates to the target location. Most would assume a different cue for each location, but this does not appear to be the case. The authors should clarify whether there was some amount of searching for the precise target location after the reach, or else how the block structure or other sensory information allowed mice to learn where exactly the target would be. In the absence of target-specific cues, some sense of how the mice achieved target-specific reach trajectories should be offered.*

> *Related to this, in Figure 1, it would be good to see some individual trajectories, as they all overlap near the target in the current plot. Clearly, the reaches were targeted, but it is unclear how targeted. Some of the adjustments at the end may reflect searching or palpation to resolve the precise spout location. It is very much ok if the mice were not reaching with micron precision each time to each of 15 different targets, but it would be good to provide the reader a better sense of what the mice were doing.*

These are important points. First, to clarify, the Cue is just a Go cue, and was the same for all targets. It is now described in the Results as "non-target-specific". For additional explanation about supplemental analyses to assess "aiming", see replies to Reviewer #1 Public Review comments above. Finally, regarding how the mice locate the target: we just don't know. As discussed above, Galiñanes and Huber found evidence for the mice using stereo sniffing, but whisking, listening to the motors, or some other strategy are also conceivable. We simply don't have data to weigh in on this. We now make this limitation clear where we describe the task.

> *In Figure 1A, CFA does not look well aligned with Tennant et al. (2011). CFA should only extend to +1 AP. The overlap of CFA and RFO seems strange. RFO also does not totally align with the injection coordinates used in An et al, biorxiv 2022.*

Thank you for your attention to these points. Our designation of the name CFA to the red dashed outline in Figure 1A was consistent with an earlier version of our previous work (Grier et al 2026) wherein we referred to the anatomical outline "MOp-ul" from Munoz-Casteneda et al 2021 as CFA. We have since revised that nomenclature to now refer to the outline as M1-fl, or the forelimb representation of primary motor cortex.

Our placement of RFO was obtained by aligning the Allen CCF from Figure 1K of An et al 2022 to our version of the Allen CCF and outlining the hotspot of RFO with a circle. We have slightly adjusted the location of RFO posterior and medial to more closely align with the injection coordinates reported in the methods of An et al. 2022 of "1.5-1.88 mm anterior from Bregma, 2.25-2.63 mm lateral from the midline." Because (as far as we understand) the injection coordinates and the map are not perfectly in register, we show a compromise between the two.

We stress that the Figure 1A map is meant to be descriptive in its illustration of the variety of organizational zones that have been identified across mouse sensorimotor cortex.

Discrepancies in the alignment procedure, animal strain, and mapping modality all introduce heterogeneity across mapping attempts that we do not aim to reconcile or resolve here.

> *Related to this, aspects of the results do seem consistent with the distinction between RFA and CFA, but this is not acknowledged or discussed. For example, the barriers in Figure 6H that lie along the M1/M2 border - these seem consistent with the gap between RFA and CFA. The same could be said for the dim trough along the M1/M2 boundary that appears to separate RFA and CFA in Figure 3B. A slightly more rostral and lateral location of CFA compared to Tennant's definition or the regions backlabeled from cervical spinal injections (see Wang, Maunze et al. J Nsci, 2018) could be expected if flattening the brain under the coverslip for imaging effectively stretches the ML axis, and Bregma (notoriously hard to define reliably at this spatial scale) was defined a bit more caudally here than in other studies. Related to this, it would be better for the field if people described their method for defining Bregma in the Methods. I suggest the authors do this here.*

We appreciate the suggestion and have acknowledged the suggested correspondence in the discussion. Given the difference in our approach from those that originally characterized RFA

(through ICMS and deep layer projection tracing) we have avoided making overly strong conclusions about this correspondence in our data. See the quoted text below.

“The spatial distribution of modulated cells in Figure 3 suggests a distinction between the caudal forelimb area (CFA, involving M1 and S1-fl) and the rostral forelimb area (RFA) in M2, while the feature gradient boundaries suggest a distinction between M1 and M2 more generally. The absence of a clearly delineated RFA was surprising, given its distinct projection patterns (Carmona et al. 2024; Hira et al. 2013b; Wang et al 2018) and functional differences from CFA (Kristl et al. 2025; Morandell and Huber 2017; Saiki-Ishikawa et al. 2025), but our results might suggest that the activity in layer 2/3 of RFA does not differ markedly from other nearby subregions of M2.”

Regarding bregma, we did not use it for atlas alignment here. Alignment was accomplished through a combination of paw vibration mapping and the location of the central sinus. Bregma’s location was only relevant for our injection of tdTomato labeling, and that labeling was used here only to stabilize the image plane. We include an estimate of it on the map solely in an attempt to be helpful, but we cannot claim we have the most reliable method for defining it.

> *The authors focus on activity aligned to cue timing. This is sensible, but it could be meaningful to know how this choice affects the definition of organization. If response clustering is largely different across time, it would seem important. I understand that addressing this question may be beyond the scope of this paper. I just wanted to raise the issue with the authors for their future consideration.*

We agree that this is important to address directly. There are two aspects to this comment: (1) does it matter if activity from approximately the same time period is aligned to the paw lift or contact instead of the cue? (2) What changes if we use data from a different period of time?

Regarding the first question (alignment), if we switch to aligning our data based on lift or contact, we have more statistically modulated neurons (see Figure 3C), but everything else is qualitatively similar with one exception: the GMM optimization doesn’t separate out the Anterior subpopulation from the Forelimb Motor subpopulation. The Anterior subpopulation only has a relatively small number of members, and they mostly exhibit the strongest peaks in their PETHs when Cue-aligned, so this makes sense. We now show the modulation maps for all of the locking events (Figure 3-figure supplement 1).

The issue of the time window is a little more complicated. There are many choices we made in this work, of course, not least of which are the task we used and the features we chose based on hand-inspection of thousands of PETHs. As we noted in the Discussion, different tasks or different features would likely distinguish more subpopulations from one another. We think of the time window as a feature choice, albeit an implicit one. We chose not to include later time points because this begins to strongly include reward signals, which are known to be large (Levy et al 2020) and can dominate other aspects of the responses. The largest differences we noted when trying time windows that extended later are that mouth-related areas are separated out in the subpopulation analyses, perhaps because of later licking/consummatory responses, but we have not explored fully enough to speak confidently on this point without much more work and another 10 figures. To keep the scope of the paper manageable, we now call out this choice explicitly (see text below). We thank the reviewer for raising these important points.

“Crafting additional PETH features, or using end-to-end neural network approaches to discover other features, might enable the discovery of additional structure (Minderer et al. 2019; Wang et al. 2023b). For example, our PETH features were chosen to be invariant to the onset time of activity, but these onset times were markedly later in lateral M1 than in adjacent M2 or S1-fl. Including onset times, using a wider window of time that includes more

of the reward/licking period, aligning data to other behavioral events, or adding other PETH features would presumably result in finer subdivisions of sensorimotor cortex."

> *The map in Figure 4 is very cool, and the spatial structure is quite striking. In terms of the actual values of the onset times in each region, I am a little concerned with a dependence on the level of reach-related activity modulation, especially relative to the level of background activity (potentially related to posture). Less reach-related activity and more background activity, which we might expect for trunk and hindlimb regions, could seemingly skew the onset times earlier. We could be getting the right answer, or an answer that makes intuitive sense, for the wrong reason. Can this potential confound be excluded with some sort of control analysis?*

The previous text wasn't clear. We have now clarified what we meant, very much in line with the reviewer's thoughts. In addition, note that our change to what is displayed in the histogram (now neurons, previously pixel values) makes clearer that there is a multi-peaked distribution of onset times and it is mostly the prevalence of each peak in each area that varies. The text now reads:

"These distributions over neurons revealed clear differences in the overall profile of activation: early onsets were more prevalent in S1 trunk and hindlimb regions, perhaps due to activity related to the animal stabilizing itself even if the neurons became more active later; then M2, and finally S1-fl and M1. Nevertheless, each area contained neurons activated at any given time in the trial."

> *The "Peak time variation" metric could potentially vary with activity level, with lower, noisier activity levels making cells appear less persistent. Perhaps a control analysis, based on SNR or some reasonable assumptions of the linkage between calcium signals and spiking, could be performed to measure the extent to which this could be creating differences between regions.*

Good idea. We have now performed this analysis, and the reviewer was correct: the correlation between peak time variation and a simple metric of SNR (assessed as range of PETH / max s.e.m.) was substantial: ρ=-0.53. We now report this correlation and describe in the Results that this metric is driven by both true peak time variation and trial-by-trial variation. Thank you for this!

"Peak time variation. To quantify whether a neuron's firing peaked at the same time for every target or varied by target, we found the peak firing rate of the response to each target, then computed the standard deviation of these peak times across targets. This value is therefore higher if the peak time varied and nearly zero if the timing was consistent. Notably, this measure correlated substantially with overall signal-to-noise ratio of a neuron's PETH (Spearman's ρ=-0.53; Methods), and thus partly measures trial-to-trial variability, not just true peak timing variability. This metric was quite low in M1, indicating highly consistent timing of the activity peak (and reliable responses), and was highest in the posteromedial part of M2 (presumably corresponding to the hindlimb representation) and the posterior tip of S1-hl (Figure 5B)."

> *One could argue that the likelihood calculations illustrated in Figure 8 are biased higher for neurons within each region since they were used for defining the likelihood for that region. I think these likelihood calculations should be done for separate neurons other than the ones used to compute the mixture model for each region.*

We agree with the point about bias: the by-area GMM in Figure 8 is biased toward cells within the area, though the effect is probably quite mild given the large numbers of neurons and modest number of parameters. However, this model was intended to make the point that even if you give an area an unfair advantage, you still can't cleanly isolate it. This was

intended to help motivate the following analysis of subpopulations, and we have now made this logic clearer. Doing it this way has the advantage that the GMM components are identical between Figures 8 and 9, while if we held out the test neurons it would not be possible to make them the same without some complicated version of bagging on the GMM components. The reviewer is right that we should make this bias explicit, though, and we have now done so:

"This mapping approach is explicitly biased toward finding feature differences between areas, allowing for a direct test of the hypothesis that response profile distributions are area-specific."

> *To me, the last Results section (Spatial overlaps between subpopulations indicate intermingled members) does two things: it shows you get the same results when you map each cell to a subpopulation independently of its area, and it shows that defining the subpopulations with cells from each area gives you essentially the same results, arguing against spatial variation of properties within subpopulations. I worry that these two points are getting merged together or not made clearly enough here, especially the first one. In general, the logic of this section does not seem well conveyed.*

Thank you for the feedback. In particular, your first point is made by Figure 9-figure supplement 4 when we fit an area-agnostic GMM to all modulated cells in the five main areas. However, your second point is one of the two main goals of the last Results section, along with the demonstration of the spatial distributions of cells after hard-clustering them by subpopulations. We have tried to clarify these main points further through substantial edits of the results section for Figure 10.

> *One set of ideas that is highly relevant and should be raised concerns an ethological organization of the motor cortex. Since the observations of Graziano, there has been a steady stream of results describing ethological organization in rodents as well. This literature is briefly reviewed in Kristl et al., Nature Communications, 2025. For example, because of the potential for a differential involvement of grasping movements across different target locations, some of the variation in neuronal tuning described in the present manuscript may stem from a region preferentially involved in grasping.*

We agree that the Graziano literature, and the substantial literature in rodent that was inspired by Graziano's work, is highly relevant to understanding the organization of motor areas. Kristl 2025 handles these issues very thoughtfully. The challenge here is that there are many possible different reconciliations of the stimulation results with ours, and some seriously unresolved challenges in doing so. To name a few:

Our subpopulations and high-gradient boundaries both give quite different pictures than microstimulation does in rodent motor and sensory cortices. In particular, microstim produces more subregions that evoke different movements than we identify, and the borders don't generally line up. This implies that the mapping between the two approaches is probably complicated.

There is a completely alternative possibility to explaining the Graziano-like results: microstimulation is thought to preferentially hit axons, and some of these projections reach the medullary motor regions. Given that the medullary motor regions have known topography in the movements they evoke (Yang et al 2023) – but may or may not be driving the movements during flexible behavior – the two approaches may not be reconcilable. Or, it may require a much deeper understanding of medulla as driving the primary movement and cortex acting as a residual controller. This is an exciting set of ideas, but as yet very underdeveloped in our understanding.

We don't know if the subpopulation structure exists at all in L5, or in the PT cells, and if it

does whether it differs. This is crucial given the frequent targeting of deep layers by ICMS stimulation protocols.

As we caution in the Discussion, it is possible that our subpopulation findings are at least partly specific to the task we used.

Although it is beyond the scope of this paper and will be addressed thoroughly in separate work, we have spent significant time with encoding models for joint angles and high-level target encoding in these same data. Given those results, we are fairly confident that the reviewer's reasonable guess, of tuning variation due to intersections between body parts, does not seem to be the main driver of the subpopulation structure we find.

After careful thought and discussion amongst the authors, we did not think that including this discussion in the paper was likely to improve interpretability of the present results for most readers. We very much agree with the point, though, and when we can narrow down the possible explanations in the future (likely in our next paper on this topic, which will address encoding) we plan to address it. We thank the reviewer for encouraging us to think through this.

> *Minor:*
>
> *(1) Page 3: "densely shared" - perhaps "broadly shared"? Dense implies most/all the neurons get the same signals, which may not be true.*

Changed to "widely".

> *(2) Page 4: "data-driven approaches" - could be more specific - isn't everything we do data-driven?*

Changed to "bottom-up".

> *(3) Page 4: "spanned areas" - perhaps "spanned multiple cortical areas", since everything spans an area.*

Changed to "spanned multiple areas" (we mention cortex just a few words earlier).

> *(4) Page 5: "intervals were generally fast" - awkward, "short" perhaps.*

Agreed, changed.

> *(5) Page 5: "which asks whether the activity for a neuron changes over time consistently in relation to any target" - Rephrase to disambiguate between consistent temporal variation in firing for all targets and variation across targets in the firing patterns. In other words, are we talking about cells that are just modulated during reaching, or cells whose firing patterns differ across targets?*

Changed ending to "to any given target". The ZETA measure really does simply ask whether there is a change in firing rate over time that is consistent across trials, for each target independently. A neuron that exhibits an identical bump for all targets would register as modulated. We chose this measure in part because of the number of temporally-modulated but untuned cells. This wasn't very clear as we had written the text, so we now note this explicitly in the Methods. Thank you for pointing out that this wasn't clear.

"For all analyses, only neurons modulated by the relevant locking event were included. Note that this measure looks for modulation over time to any target; it is indifferent to whether the neuron exhibits tuning across targets."

> *(6) Figure 1: It seems like some of the abbreviations used in 1A have not been defined yet in the paper.*

Yes. It's a long list, and we wanted to put the citations for the description of each area together with the definition of the acronym. Moreover, we wanted all this info together with the description of how we aligned these area descriptions from others' work with one another on the Allen atlas. This was impractical in the caption, and would be a long digression for what is intended as a simple point in the Results, which is why we refer to the Methods here.

> *(7) Page 8: "Given that these areas have known spatial organization within them and structure was apparent by eye in the spatial scatterplot of modulated neurons (Fig. 3A)," - it is not clear what spatial structure we are supposed to see in 3A.*

Good point. We have changed the parenthetical to: "(for example, the less modulated band along the M1/M2 border in Fig. 3A)"

> *(8) Page 8-10: The region-wide onset analysis breaks up the flow from PETHs to the metrics used to quantify them. I suggest moving this section (Onset of neural activity varied with somatotopy and subregion) to later in the manuscript.*

We appreciate the reviewer's input on organization. We went back and forth many times in how to organize the many results in this paper. The reviewer is right that this analysis breaks the flow, but the reason we included it where we did was threefold. First, it uses an easily-understood metric to introduce the reader to how we made maps from single-neuron features. Second, it easily introduces the power of making such maps. Finally, it makes clear that if we are not careful with how we handle time in the feature design, timing will dominate.

All these things said, this has helped inspire us to add a result in which we re-examine timing broken down by subpopulation (Figure 9-figure supplement 2C). It shows that subpopulations timing distributions appear more distinct than distributions for areas, but there is still substantial heterogeneity in timing that is explained by location in cortex and not subpopulation membership alone.

> *(9) Page 12, Target tuning linearity: This metric should be clarified in the Results. It is not clear how the 2D of targets is turned into 1D. Also, the plot in the figure has correlation on the y-axis, and it is not clear how each target location gets its own correlation value. The phrase "optimized anchor target" is unclear.*

Agreed this needed to be clearer. The text in the Results now reads: "To quantify how linearly a neuron's activity related to target location in physical space, we correlated the 15D vector of mean activity of the neuron for each target with the 15D vector of the targets' ordinal distances from the neuron's preferred target (Methods)." In agreement with your suggestion, we have dropped use of the phrase "anchor target" in favor of "preferred target", which should be clearer. We have also revised the Methods text accordingly to clarify.

To directly answer your question, we turn the targets from 3D positions into 1D by computing the ordinal distance of each target from a preferred target. (Note that the preferred target is actually the one that maximizes the resulting correlation; this is detailed in the Methods). There therefore aren't 15 correlations; we're correlating two 15D vectors, where each has one element per target and the "ordinal distance" vector has a zero for the preferred target. Hopefully the new description makes this clearer.

The figure schematic was unclear, thank you for catching that. We have updated the Y axis to read "mean activity" and the X axis now reads "dist. to pref. target."

> *(10) Page 12, paragraph beginning "We also compared our metric maps simply using the top 20 PCs." - This paragraph is unclear, since both sentences refer to using the metrics. I would guess the authors mean that the metric maps were compared with and without PCA and basis rotation, but this is not clearly stated.*

Thank you, this was unclear as written. We have changed it to:

"We also compared our metric maps with maps generated from the top 20 PCs of the PETHs (Methods), rotated using VARIMAX to identify a sparser basis (Musall et al. 2019)."

> *(11) Page 18: "These results make clear that the working hypothesis - of areas with well-separated feature distributions - is incorrect." This is the clearest statement of the impact of the results. The authors could consider including this in the Abstract or Introduction.*

Thank you for pointing this out. We agree, and have added a similar phrase to the Abstract.

> *(12) Figure 9: It would be great to also just see the average PETHs for each of the four clusters to get a better sense of how their time series differ.*

Good idea. The feature computations are a many-to-one mapping, so it's not possible to literally generate a PETH from the mean of the cluster, but we have added PETHs from well-modulated neurons that are near the means of their subpopulations (Figure 9-figure supplement 1).

> *(13) Figure 9B: Colorbar has no label.*

Fixed, thanks.

> *(14) Figure 9C: Need a colorbar - need to see the difference in density for locations.*

The color map is the same Figure 8B, which is now noted in the caption for Figure 9C. The scaling of likelihoods is almost totally uninformative; they're not well-behaved like probability distributions, so you'll note that even on Figure 8B the labels are simply "max likelihood" and "min likelihood". The important pieces of information here are that these are log likelihoods (noted in the Figure 8 caption), and the visualization of the color map itself (from the color bar). Given these considerations, we have elected to keep the maps themselves a little larger by not trying to squeeze in a minimally-informative colorbar to all of the plots, but thank you for noting that the reference to 8B was needed.

> *(15) Page 22: "additional spatial structure could be present" - The nature of the additional spatial structure here is a bit opaque. The authors could clarify what additional structure may be present.*

Good idea. This paragraph now reads:

"The overlaps in the subpopulation likelihood maps above imply that members of different subpopulations are spatially intermingled, but it is less clear whether each subpopulation has homogeneous response profiles across space. In particular, the use of likelihoods mixes two properties: the fraction of neurons in a given neighborhood that are members of each subpopulation, and the heterogeneity of response profiles amongst members of that subpopulation. These properties could vary systematically with respect to one another, and the spatial structure shown by the likelihood map does not disentangle them."

> *(16) Figure 10E, legend: "GMM component" - I think this should be "GMM subpopulation" to avoid confusion with the previous use of "component" above, referring to the components of the GMM models for each region.*

Thank you – good catch. Changed to "Likelihood map".

> *(17) Page 24: "Note that this consistency also validates the use of clustering to combine components and identify the subpopulations in the first place." - I don't totally get this, and how this result validates the method of combining components, as opposed to just clustering all the cells from all regions at once. Perhaps the implied opposing strategy is not clear here.*

We have changed this sentence to:

“Note that this consistency mirrors the low Bhattcharyya distances between corresponding GMM components in Figure 9B, and further validates the use of clustering to combine components from different areas.”

Regarding the reviewer’s larger point, we have three thoughts. First, we do also show the result of fitting the GMM to all cells together (Figure 9-figure supplement 4).The result is similar, but the Anterior subpopulation is lost because its membership is low and so the ICL criterion can’t justify a fourth cluster. Second, because we imaged more neurons in some areas than others, fitting the GMMs to each area separately put their representations on a more equal footing. Finally, doing the analysis this way allowed us to most directly compare our two hypotheses, as illustrated in Figures 8A and 9A.

> *(18) Page 25: "in the zones where different subpopulations overlapped" - I would omit this, since "intermingled" seems to mean exactly this.*

We included this phrase to prevent quickly-skimming readers from incorrectly concluding that the subpopulations overlapped entirely and were therefore intermingled everywhere. The reviewer is right that it’s unnecessary for a careful reader, but we aimed to prevent misinterpretation by readers that might skip to the Discussion for a results summary.

> *(19) Page 25: "content of the activity, but also its format" - the difference between content and format is not entirely clear. Metaphor not quite metaphoring here. Agreed. We have added examples to clarify.*

“This makes clear that there are potentially important differences not just in the content of the activity (e.g., encoding target vs. movement commands (Grier et al. 2026)), but also its format (e.g., linear encoding vs. nonlinear, persistent vs. brief responses).”

> *(20) Page 30, bottom: In the description of the behavior, more details should be provided, especially since the paradigm is new. For example, it says the block size was reduced - what was the ultimate block size?*

Targets were cued randomly in the behavior performed during neural recordings. Blocked trials were used during training and were phased out incrementally as performance improved. This and various other details have been added. Please let us know if there are other specific details you would like to see in the final version.

> *(21) Page 39, citation of An, Mulcahey et al.: There is a biorxiv version with a different author list that could be cited.*

This was an error with our citation manager, and has been corrected. Thanks for catching it.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *Overall, this is a remarkable study with well-designed in-depth analyses, and I only have some minor suggestions that could help improve the clarity of the paper.*

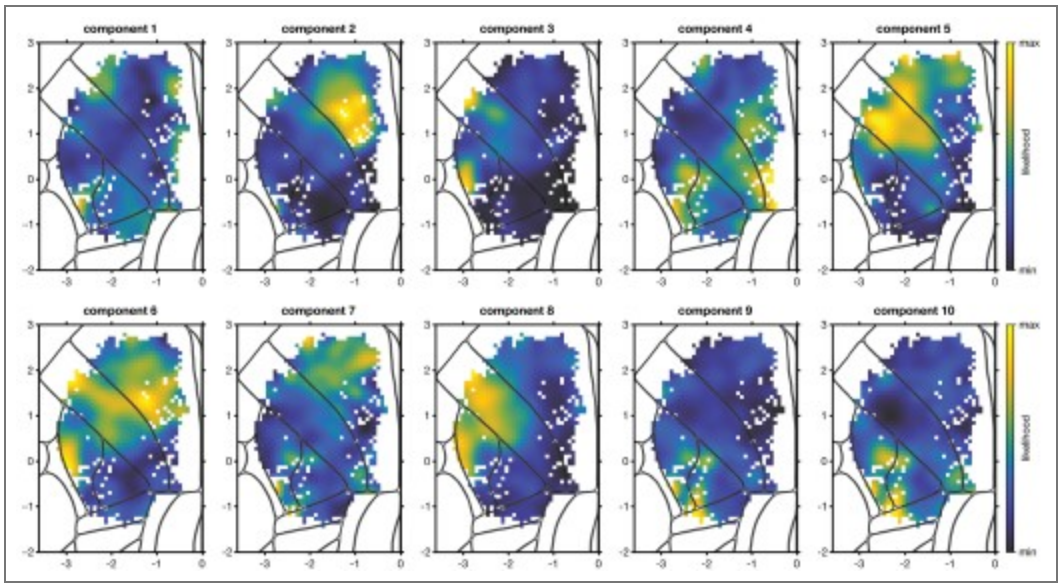
Thank you!

> *General:*

> *It is not immediately clear to me why the GMM approach used in this study is more interesting than a clustering approach based on single-neuron response patterns (See Esmaeli et al., Neuron 2021 or Oryshchuk et al., Cell Report 2024). But my impression is that it led to the same observation that most clusters are widely distributed across cortical areas, with different proportions, but a few clusters are quite specific to a few areas. A noticeable difference perhaps is the number of clusters - or response profile - that seems particularly low (only 4) in the current study. Could the authors clarify and comment on that, maybe?*

The reviewer brings up an interesting point: at heart, these works ask related questions, albeit about different effectors, tasks, recording modalities, and types of information encoded. Those differences probably mean that results cannot be directly compared, but we can certainly discuss the methodological tradeoffs. The two papers mentioned take a more traditional first step, using PCA on the vectorized PETHs to reduce dimensionality, then layer on a spectral approach to improve clusterability. These are good methods; we use something similar as our alternate method, applying VARIMAX to the PCs instead of spectral methods to preserve linearity of transforms. For the kinds of responses both they and we have, PCA will tend to most strongly pick up two aspects of the responses: tuning and timing. This is because vectorized PETHs will have large values in the rows corresponding to the target/condition and time points where the high activity is, and the alignment of these profiles with those of the other neurons will capture a large fraction of the variance. For data like either theirs or ours, this would tend to cluster apart left-tuned cells from right-tuned, and (more importantly here for revealing spatial structure) early-response cells from later response cells. That intuition is consistent with what those papers report, and examining our VARIMAX'ed PC plots closely (which have sharpened in the latest version thanks to improved normalization), we can see that they break apart sub-regions largely based on timing. In our feature approach, we intentionally chose our features to be largely invariant to both tuning preferences and timing. Instead, we chose our features to pick up on what we call the single cell "response format": response duration; peak time variation (but not absolute timing); and tuning sharpness, persistence, and linearity. These different methods pick up on different aspects of responses.

To double-check that the PCA-then-spectral approach reveals similar structure to our use of VARIMAX on the PCs, we tried applying the suggested method to our data. We applied spectral clustering to the N x 20 PETH PC feature matrix, then fit an area-agnostic GMM to the spectral features. We plot the likelihood map for the components of a GMM with 10 modes. The GMM components did not display clear spatial structure beyond that observed in the VARIMAX'ed PCs (Figure 5-figure supplement 1) and were less interpretable than those identified by area-agnostic clustering of our response features (Author response image 1). As noted, the number of subpopulations identified by the clustering of our hand-engineered features is lower than what would be obtained from clustering the PCs of the PETHs. This is likely the result of the substantial heterogeneity in activity onset and preferred target that is preserved by PCA. Because our central approach is largely agnostic to these two sources of variation, the number of identified clusters reflects the dominant patterns of variation beyond these two sources.



Author response image 1. GMM fit to spectrally transformed PETH PCs, agnostic to anatomical areas. One GMM was fit to the spectrally-embedded PC feature vectors of cells from all 5 main areas. Each component of a 10 component model is shown.

> *Also, I think it would greatly help the reader to return to PETHs at some point, if possible, to show the response profiles of each identified neuronal subgroup (page 20). To what extent are they similar or different across the cortical areas (for the same neuronal subgroup)?*

This is a good idea. We have added a figure to address this question and the related question by R1 (Figure 9-figure supplement 1). In short, given the wide variety of PETHs we observed, there is of course still substantial variation within subpopulation, and some mild but systematic differences in the distribution of what we observe across areas. We now discuss the conclusions from this plot in the Results:

"As a qualitative depiction of the response profiles identified with each subpopulation, we plotted the two highest-likelihood cells for each area/subpopulation combination (Figure 9-figure supplement 1). These examples reveal stereotypy in the subpopulation responses across areas, but also show variation across areas, especially for the two somatomotor subpopulations."

> *Specific:*
>
> *(1) Figure 2B and M&M: the 3D spatial organization of the target locations is not immediately clear. What is the spacing between target locations? What is the 'final azimuthal spacing'?*

Added, thanks. The pairwise horizontal distances between targets were between 1.72 and 6 mm apart and the vertical spacing within a column was 1 mm. "Final azimuthal spacing" just referred to the targets being closer together during training and our gradually spacing them apart to their final locations. We have also added some relevant details about the training.

> *(2) Figure 2C: It would help to have a scale bar (mm).*

Added, thanks.

> *(3) Figure 2C: It would be easier to appreciate the variability of the trajectories across trials to plot an overlay of trajectories to one target only (could be a Supplementary Figure).*

The reviewer has a good point: the variability and accuracy of aiming was hard to ascertain from the plot. We experimented with a few options for making this clearer most effectively. We have now added Figure 2-figure supplement 1 that shows in the third subpanel of panel A the finger centroid trajectories for one of the 15 targets highlighted for the mouse shown in Figure 2C, mouse 3. The centroid trajectories for all other mice are shown as well to illustrate similarities and differences across animals as well as the overall variability. As noted elsewhere we have also included an analysis of the variability of the centroid trajectories, showing that reaches to a given target were more similar than reaches to different targets. We think this provides a fuller picture of the behavior and intend to provide still more detail in future work. Thank you for suggesting additional detail here!

> *(4) Figure 4: It would be nice to also show the amplitude-normalized grand-average PETHs for the different areas.*

This is an interesting suggestion. After careful consideration, we think that this analysis is not as effective for depicting overall timing and modulation profiles as the current ones, given the strong amount of target selectivity and response time heterogeneity (now better visible in the revised Figure 4A). When computing the grand mean of all cells within each area, the dominant features distinguishing areas are onset time and response duration. The differences across areas in these two features are better supported by the analyses of Figures 4 and 5 due to the large amount of heterogeneity in responses within each area. We thank the reviewer for encouraging this exploration; more complicated spin-offs will likely inform additional timing analysis in the next paper on these data.

> *(5) Figure 7C: figure legend - although it is quite self-explanatory, please explicitly indicate which pattern corresponds to the 'Three contour levels (98%, 95%, 90%)'.*

We have now added this as a legend on the figure panel itself (here and on similar plots). Thanks for pointing this out.

> *(6) Figure 8: Is there also an interesting asymmetry between sensory are motor areas, with neurons in sensory areas being more likely associated with motor areas (B and C), whereas neurons in motor regions are less likely to arise from the distribution of sensory areas (dark blue color in frontal regions in D, E, and F)?*

This is an interesting observation, but we understand it to be an artifact of colormap scaling. As mentioned above, likelihoods are not well-behaved like probability distributions are: for example, they are not bounded at 1, and their sums over a dataset can have any positive value. The only things that can be interpreted are their relative values. This makes their scaling functionally arbitrary – you'll notice we used "min likelihood" and "max likelihood" instead of numbers, which would be nearly meaningless – and therefore presents a problem for scaling the colormaps. We don't know of a principled way around this problem. To deal with it, we simply put the ends of our colormap at the extreme pixel values. It so happens that both the M1 and M2 maps had a handful of neurons in a less-sampled spot at the bottom of M2 that were very low-likelihood, which results in what you noticed. We debated removing those neurons for this purpose, but we had no basis on which to do that kind of manipulation, so we left it as the most honest representation of the data we could produce.

To clarify this, we now mention in the caption "The ends of the colormap were set to the maximum and minimum likelihood values for each map."

> *(7) Figure 9B: there are two-time 'S1-hl: 1' indicated at the two bottom rows of the distance matrix. I suppose one of them should be 'S1-tr: 1' instead?*

Fixed, thanks for catching it.

> *(8) Page 20: 'This hinted at a second hypothesis: that some of the 'modes' (groups of neurons) discovered separately in each area might correspond.' ???*

We had meant “mode” as in “multimodal”, but it was very unclear. We have rewritten the sentence:

“This hinted at a second hypothesis: that a peak in the multimodal distribution from one area might correspond to a peak in the multimodal distribution of a different area.”

> *(9) Figure 9S2: Please indicate for which area each map is computed.*

The caption was not clear enough about what we were doing here: we fit the GMM on all neurons together, ignoring which area they came from. We have now clarified it in the caption:

“One GMM was fit to the feature vectors of cells from all 5 main areas. Each map plots the likelihood for all cells to each of the three components of this area-agnostic GMM.”

> *(10) M&M, Subjects and surgical procedures: 'ambient temperature of 71.5 {degree sign}F', please use international units.*

Done.

https://doi.org/10.7554/eLife.109240.2.sa0