The natural variability of a dexterous motor skill is stably encoded in the cortex of freely behaving mice

Elizabeth A. de Laittre[1] and Jason N. MacLean[2]
[1] Committee on Computational Neuroscience, University of Chicago, Chicago, IL, USA
[2] Department of Neurobiology, University of Chicago, Chicago, IL, USA
Correspondence should be addressed to: jmaclean@uchicago.edu

The nervous system enables precise, skilled movements across diverse contexts, consistently performed day after day. However, the extent to which the brain's neural activity patterns are specific to each instance of a movement or behavior, and how similar these patterns are across days, remains unclear. To address this, we record calcium fluorescence activity in the motor cortex of freely moving mice performing a self-initiated, skilled reach-to-grasp task. The high trial counts and single-trial variability in this task allow for a rigorous statistical analysis of moment-to-moment movement encoding across matched behavioral sets over five days. We found that single neurons in motor cortex encode the single trial (reach-to-reach) details of paw, digit, and head movements, and this encoding is stable over days. This suggests that the stable contribution of individual cells in the motor circuit underlies the generation of skilled movements, even in complex sensory-driven tasks like reach-to-grasp.

Introduction

Even well-practiced behaviors exhibit variability with each attempt. In basketball, for instance, no two shots are identical, as each is shaped by factors such as starting posture, fatigue, and interactions with defenders and teammates. Yet, players consistently adapt and successfully perform this precise skill across these dynamic contexts. While generating goal-directed movements under such conditions remains a significant challenge in robotics (Billard and Kragic, 2019), biological nervous systems excel at this. To understand how motor skills are consistently executed across varying contexts, it is crucial to investigate how movement variability is represented in neural circuits and whether the representation of learned, complex motor behaviors remains stable over time. This insight can shed light on the circuit-level precision required to adapt to perturbations while maintaining reliable motor performance over extended periods of time.

Prehension, the act of reaching and grasping objects, is a vital yet complex movement requiring precise coordination of the trunk, shoulder, arm, wrist, and fingers, guided by rapid sensory feedback processing. The need to interact flexibly with diverse objects further enhances its complexity. Recent work suggests that precision reaching in mice relies on online sensory feedback rather than pre-planned, ballistic movements (Becker et al., 2020). Motor cortex plays a central role in the brain-wide circuit controlling reach-to-grasp and other precise reaching behaviors in rodents, as shown by lesion and inactivation studies (Guo et al., 2015; Wang et al., 2017; Miri et al., 2017; Galiñanes et al., 2018). In mice, the caudal forelimb area (CFA) and other cortical motor regions represent high-level categorical variables, such as behavioral category (Dhawale et al., 2017; Peters et al., 2014, 2017; Huber et al., 2012; Dombeck et al., 2009) and reach direction (Galiñanes et al., 2018). Simultaneously, motor cortical activity captures low-level features of forelimb movements, including muscle activations and kinematics (e.g., position and velocity), during ladder walking (Omlor et al., 2018), climbing (Koh et al., 2024), and a precision reach-for-water task (Zhu et al., 2022). However, the extent to which CFA in mice encodes whole-body posture or digit movement details during sensory-driven reach-to-grasp behavior in a freely moving animal, which requires coordination of the whole body, from gross movements of all limbs and

the trunk to fine control of the digits, is unclear. While there is evidence for coordinated digit use in mice (Whishaw et al., 2017; Barrett et al., 2020; An et al., 2024) and posture-related activity in rats' posterior parietal, secondary motor, and sensorimotor cortices (Mimica et al., 2018; Disse et al., 2023), it is unknown how these kinematic features are represented in CFA. Understanding how CFA encodes movement features—such as proximal forelimb actions, whole-body posture, and fine digit control—is essential for clarifying its role in complex behaviors like reach-to-grasp. If M1 contributes to the online control of these movements as part of a broader sensorimotor circuit, it is likely to encode their moment-to-moment details.

Neural encoding in sensory and association brain areas often shows session-to-session variability, a phenomenon termed representational drift (reviewed in Micou and O'Leary, 2023; Driscoll et al., 2022; Bauer and Rose, 2021; Chambers and Rumpel, 2017; Clopath et al., 2017; Lütcke et al., 2013). In contrast, motor areas exhibit more stable encoding over time (Jensen et al., 2022; Dhawale et al., 2017; Katlowitz et al., 2018; Liberti et al., 2016; Chestek et al., 2007; Fraser and Schwartz, 2012; Ganguly and Carmena, 2009; Gallego et al., 2020; Flint et al., 2016; but see Rokni et al., 2007 and Liberti et al., 2016). Most studies of the stability of motor encoding have relied on trial-averaged metrics, such as peri-event time histograms (PETHs), which can mask neural stability or instability by not accounting for behavioral variability within and across sessions. This raises the possibility that apparent neural instability could reflect changes in behavior rather than shifts in the mapping between neural activity and behavior (Liberti, Schmid, et al., 2022; Jensen et al., 2022).

The stability of moment-to-moment encoding for fine-grained movement features, such as joint position, velocity, or muscle activations, remains underexplored. This is partly due to the challenge of studying behaviors that are variable enough to reveal specific neural activity patterns for each instance of a movement, while maintaining sufficient control to ensure statistical rigor. To address this gap, we examine a complex whole-body, dexterous reach-to-grasp task that produces high trial counts with significant trial-to-trial variability within sessions. Despite this variability, individual animals consistently explore a similar kinematic movement space across sessions, enabling us to create matched datasets for directly comparing the encoding of movement variables—kinematics of the head, paw, and digits—across days.

Results

A dexterous reach-to-grasp task elicits variable yet spatiotemporally coordinated forelimb and digit movements. To investigate the stability of the cortical encoding of moment to moment forelimb and head movements, we trained unrestrained mice on a challenging reach-to-grasp task. We trained the mice to reach for a millet seed placed on a narrow pedestal positioned in one of three locations just outside the front of the arena (Fig 1A, B, C). Mice self-initiate the delivery of a seed from the automated system by breaking an infrared beam positioned at the back of the arena opposite the reach port (~15cm or 2 body lengths away). Due to the limited mobility of the mouse shoulder, a critical aspect of the task is positioning the body and head to ensure the pedestal is within a reachable angle through the reach port. After triggering the delivery of a new seed and returning to the reach port, mice performed a series of 1 to 10 individual reaches in quick succession until they either successfully retrieved the seed or knocked it off. Each reach is defined as a single out-and-back motion of the paw through the reach port, where the paw exits and re-enters the arena. A set of reaches performed in quick succession before moving to trigger the delivery of another seed is referred to as a bout. Throughout this paper, we treat each reach as a single trial. During a 40 minute session, a trained animal will perform anywhere from 789 to 2831 reaches and successfully retrieve and consume anywhere from 27 to 139 seeds (see Supp Fig 1).

The pedestal's diameter is similar to that of the seed, making it easy for the millet seed to be knocked

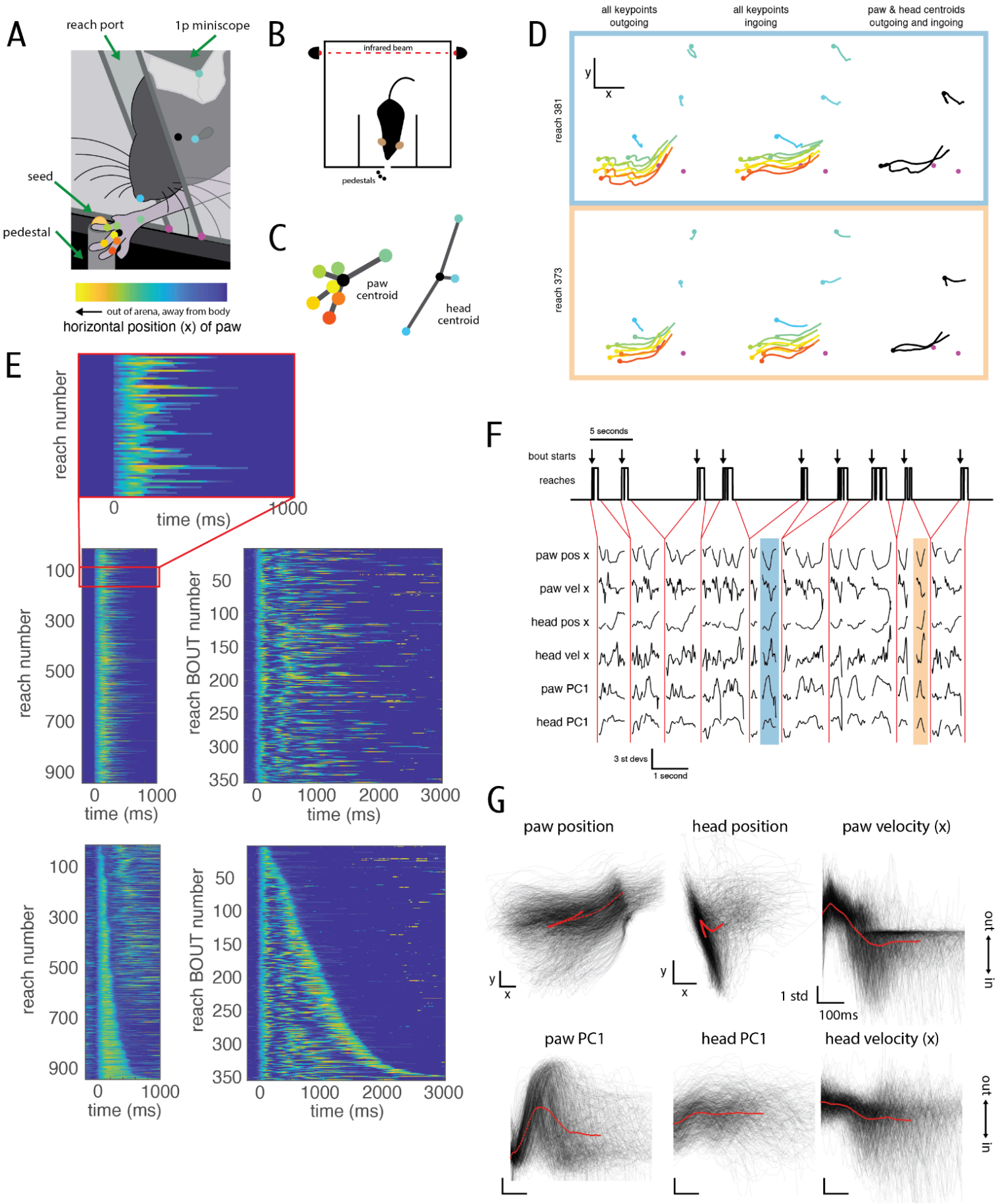
A
reach port
1p miniscope
seed
pedestal
out of arena, away from body
horizontal position (x) of paw
B
infrared beam
pedestals
C
paw centroid
head centroid
D
all keypoints outgoing
all keypoints ingoing
paw & head centroids outgoing and ingoing
reach 381
reach 373
y
x
E
reach number
time (ms)
reach BOUT number
F
5 seconds
bout starts
reaches
paw pos x
paw vel x
head pos x
head vel x
paw PC1
head PC1
3 st devs
1 second
G
paw position
head position
paw velocity (x)
paw PC1
head PC1
head velocity (x)
1 std
100ms
out
in

Figure 1: A dexterous reach-to-grasp task elicits variable yet spatiotemporally coordinated forelimb and digit movements.

A) Closeup view of mouse reaching, with 10 keypoints on the paw and head marked, from the viewing angle of the tracking camera. Colorbar at bottom indicates paw position in x relative to the view shown here; relevant for the heatmaps in E.

B) Schematic of mouse in arena, indicating location of pedestals, reachport, and infrared beam

C) Closeup on keypoints on the paw and head, and their distances to the paw and head centroids (black dots). Colors match those in A.

D) Two example reconstructed reaches, where the colors match those in A and C. For all panels, dots indicate the endpoint (point of the reach that is furthest from the arena). Left column, all paw and head keypoints during the outgoing part of the two reaches. Middle column, same for the ingoing part of the reaches. Right column, the paw and head centroids only for both the outgoing and ingoing parts of the reaches. Magenta dots indicate the base corners of the reach port.

E) Paw position in the X dimension plotted for all reaches and bouts in one session for an example mouse (mouse 5). Colormap is under the schematic in A; yellow colors indicate further outside the arena, whereas blue colors indicate closer to or inside the arena. Top, reaches sorted in the order in which they occurred (left), and the same reaches resorted into bouts, again in the order in which they occurred (right). Bottom, the same reach and bout data sorted according to reach and bout duration.

F) Top, binary variable indicating when the animal is reaching (paw is outside the arena), with arrows at the start of bouts. Bottom, time courses of 6 of the 14 kinematic variables extracted via keypoint tracking, during the bouts plotted above.

G) All reaches in one session for mouse 5, for the same 6 variables as in G. Individual reaches are plotted in transparent gray, the average (time-locked to reach onset) is plotted on top in red. Scale bar for paw and head position correspond to 30 pixels in x and y; the view is the same as is schematized in A and for example reaches in D. Scale bar for paw velocity, head velocity, paw PC1, and head PC1 are 100ms and 1 standard deviation of each variable.

off (Fig. 1B). Successfully retrieving the seed requires precise paw placement and highly coordinated, precisely timed digit control. The narrow pedestal prevents repositioning the grasp on the seed after initial contact, allowing the mouse only one attempt per contact. After grasping the seed, mice must maintain their grip while transporting it back into the arena, adding to the task's complexity. Performance on this task typically plateaus after about 10–14 days of training (Chen et al., 2014).

To quantitatively describe the animals' digits, paw, and head movements during reaching, we recorded high-speed video during all sessions and employed markerless keypoint tracking (DeepLabCut, Mathis et al., 2018; Nath et al., 2019) to reconstruct the kinematics of each reach (see Fig 1A, C, D). On the paw, we tracked the joints of the first three digits and the wrist; on the head, we tracked the nose tip, corner of the eye, and a prominent corner of the surgical implant (Fig 1A). From this keypoint tracking, we extracted the position and velocity of the paw and head centroids during reaching (Fig 1C); we also ran principle component analysis on the collection of distances of the paw and head keypoints from their respective centroids, to extract proxies for the shape and orientation of the paw and the orientation of the head (see Methods). The position and velocity of the paw centroid indicate activation of proximal forelimb muscles controlling the elbow and shoulder, while principal components of the conformation of the paw reflect the activation of distal muscles controlling the wrist and digits. Similarly, the position and orientation of the head reflect the whole-body posture of the animal and potentially the activation of shoulder and neck muscles.

Within a single 40 minute session, animals exhibit high reach-to-reach variability in the speed, shape, and duration of their digit, paw, and head movements during reaching (Figs 1E, F, G). This trial-to-trial variability persists even after the animals have been trained beyond the time needed to master the task, as evidenced by their stable success rates during the period of study. (Supplementary Fig 1). We considered this within-session variability advantageous, as it allows us to examine the moment-to-moment neural encoding of these movements (as in Omlor et al., 2018 and Becker et al., 2020).

Each animal develops its own unique reaching pattern, yet consistently uses a similar range of movements across days. To investigate whether the neural coding for this behavior changes or remains consistent over time, we analyzed data from sessions on five consecutive days late in training (starting on or after day 20), a period during which animals demonstrate stable success rates (Supplementary Fig. 1, success rates across 5 days). The within-session variability observed on a given day persists across the five days of study, with the variability centered around very similar means (two example animals in Fig 2A). The specific digit, paw, and head movements employed differs across animals, as each develops its own strategy for retrieving the seed (see Supplementary Fig. 2 for kinematic descriptions across all five mice). In some cases, the strategy involves rhythmic, consistent reaches, with individual reaches limited to a narrower range of magnitude and duration. In other cases, there is greater variability in reach magnitude, duration, and timing within bouts (see the temporal patterning of the two example mice in Fig 2B) For each animal, the distributions of all measured kinematic properties and the corresponding success rates remain consistent across all five days. This stability is further validated by the covariate balancing procedure used for the linear models in later sections: after matching, the majority of reaching timepoints are retained across all days, indicating a similar joint distribution of kinematic variables over time (Supplementary Fig. 3; see Methods). This consistency in movement across sessions enables the use of this dataset to compare neural encoding of reach, grasp, and carry behaviors over multiple days.

Neurons in CFA M1 motor cortex modulate during reaching. We recorded the activity of populations of individual neurons in caudal forelimb area of mouse motor cortex (CFA, also referred to as forelimb M1) using an Inscopix miniscope to perform one photon calcium imaging in freely moving (non-headfixed) mice. We combined the miniscope with a prism to achieve an imaging plane perpendicular to the cortical surface (Andermann et al., 2013; Resendez et al., 2016). This setup enabled simultaneous imaging of putative cortical layers 2/3 and 5a, capturing hundreds of cells in a single field of view (244 +/- 50 neurons [138 347] (mean +/- st dev, range) per mouse per day). By consistently returning to the same field of view each day, we were able to identify and longitudinally track the activity of many of the same cells over multiple days.

The majority of neurons exhibited significant modulation during reaching, as determined by an ANOVA comparing neural activity during reaching with that during other behaviors, including locomotion, grooming, food manipulation, and food consumption (5544 out of 6105 *cell-days* (90.8%), where each cell-day represents a single day's observation for a given neuron; $p = 0.05$ with Bonferroni correction for the number of cells in each field of view for that day; see Methods). Across the recorded population, neurons displayed diverse temporal response profiles, with activity observed before, during, and/or after reaching (Fig. 3C). Additionally, cells were active during the intervals between reach bouts, when the animal moved to the back of the arena to trigger the next seed delivery. In this analysis, each reach was treated as a trial. Neural activity showed notable variability across individual trials (Fig. 3C, D), reflecting the variability in the animals' movements. Trials are sorted on each day by reach duration, as in Fig 1E bottom left. Not all neurons were active during every reach, and even when a neuron was active, its activity magnitude varied significantly. Despite this variability, neurons demonstrated stable firing patterns across days (Fig. 3D), both in single-trial activity (middle) and in their trial-averaged response (bottom).

Reach-locked average neuronal activity patterns (PETHs) are stable across days. We quantified the average response of each neuron using peri-event time histograms (PETHs) time-locked to the start of the reach. The variety of response timecourses observed in single trials (Fig. 3) was also evident in the PETHs (Fig. 4A). While many neurons showed peak activity during reaches, others were most active

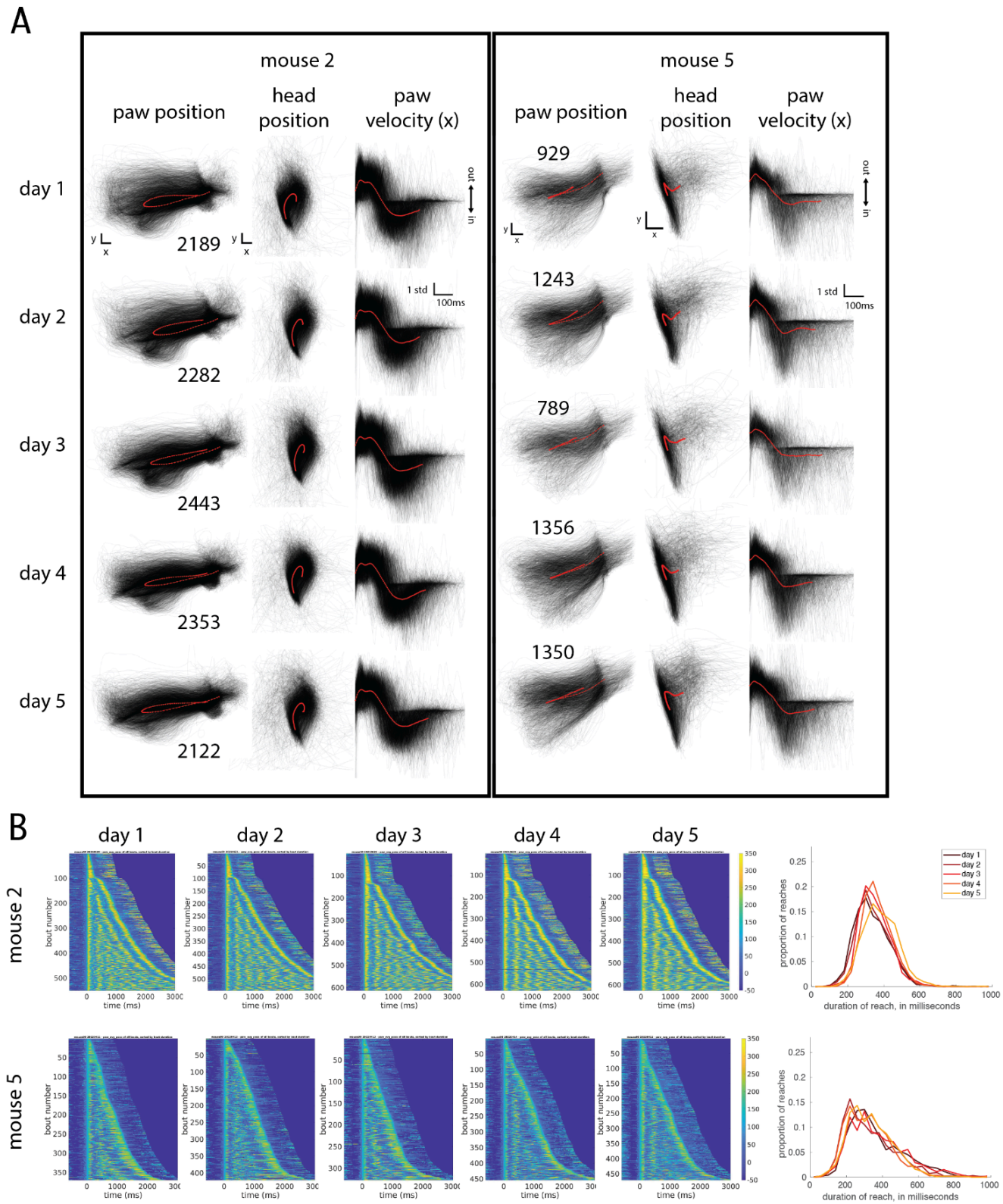

Figure 2: Each animal develops its own unique reaching pattern, yet consistently uses a similar range of movements across days.

A) All reaches in each of five sessions for two example mice, for 3 of the 14 kinematic variables extracted via keypoint tracking. As in Fig 1G, individual reaches are plotted in transparent gray, the average (time-locked to reach onset) is plotted on top in red. Numbers next to the paw position plots are the number of reaches in that session.

B) For the same two example mice (mouse 2 and 5), for all five days: Left, paw position in the X dimension plotted the same way as in Fig. 1E. Right, histograms of duration of individual reaches.

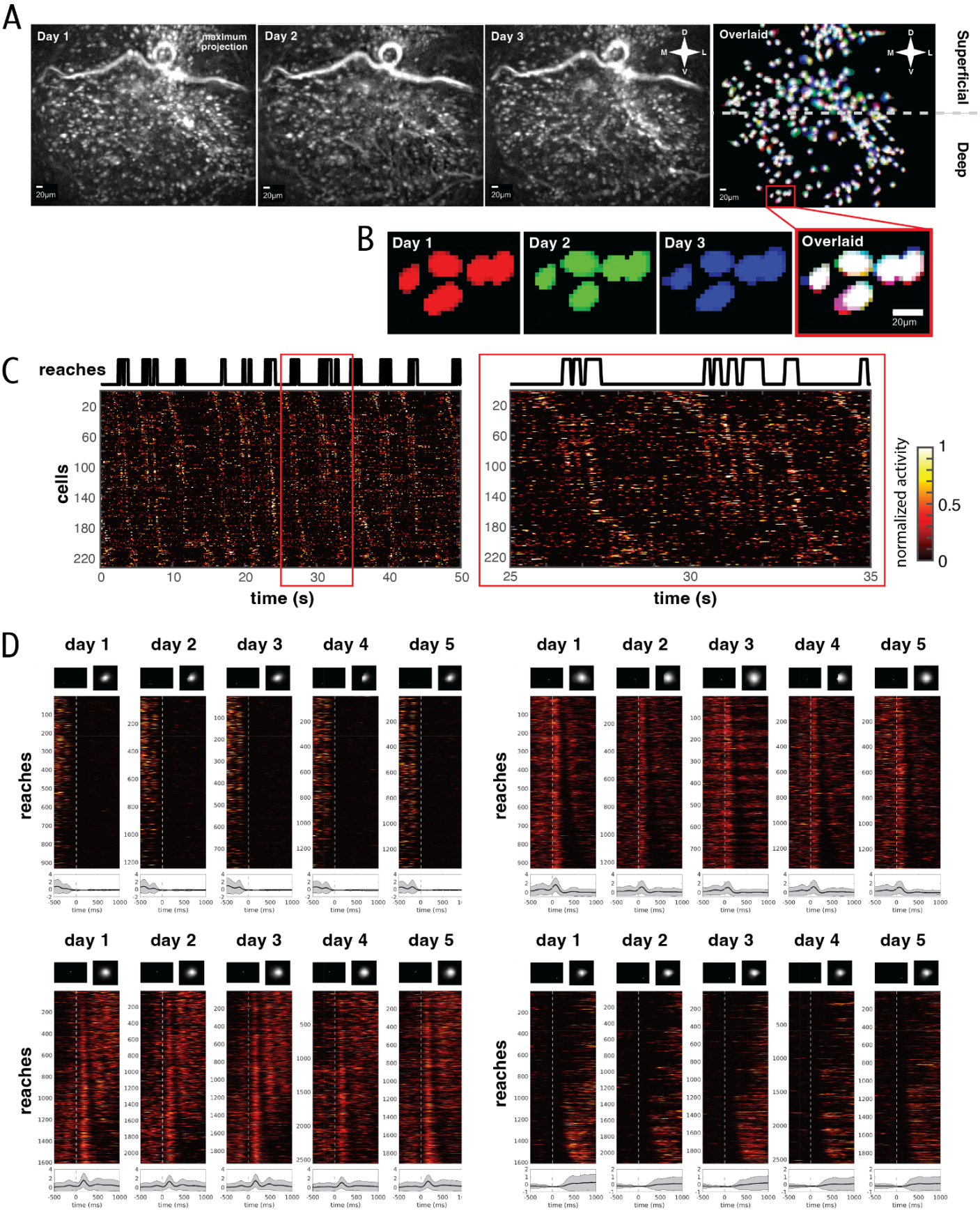
A
Day 1
maximum projection
Day 2
Day 3
Overlaid
Superficial
Deep
20μm
B
Day 1
Day 2
Day 3
Overlaid
20μm
C
reaches
cells
time (s)
normalized activity
D
day 1
day 2
day 3
day 4
day 5
reaches
time (ms)

Figure 3: Neurons in CFA M1 motor cortex modulate during reaching

A) Far left, middle left, middle right: Maximum projections of the motion-corrected calcium imaging movie of the same field of view (mouse 4) on three subsequent days. Compass indicates dorsal-ventral and medial-lateral. Far right: overlay of footprints of individual cells identified on all three days, with day 1 in the red channel, day 2 green, and day 3 blue.

B) Same as in A, but zoomed in on five cells found on all three days.

C) Top: Binary variable indicating when the mouse was reaching. Bottom: Heatmap of inferred spiking activity of all recorded cells from one field of view for 50 seconds of recording. Right inset: Temporal zoom to 10 seconds of recording. Each cell's activity has been normalized separately within this window: minimum value = 0, mean + 8*standard deviation = 1.

D) Single-trial inferred spiking activity of four example cells found on all five days. For each day, the top panel shows the spatial footprint of the cell (right) and its location within the field of view (left). The middle panel shows a heatmap of the single trial activity. Normalization is the same as in C. Trials are sorted according to the duration of reaches, as for the kinematics in Fig 1E: short reaches at the top, long reaches at the bottom. The bottom panel shows the average response across all trials on that day (PETH) as mean +/- st dev.

either before or after the reach. This suggests that these neurons may be modulated by other task components associated with the reach behavior, such as positioning the body at the reach port, locomotion to trigger the next seed delivery, or consuming seeds.

We found that, across the population, PETHs were highly similar across days. This is demonstrated visually by showing the PETHs for all cells tracked across all five days for an example animal (Fig 4A), sorted according to peak activity modulation on either day 1 or day 5. Regardless of sorting day, the rank order is maintained across all other days. To quantify similarity across days, we computed the correlation between each longitudinally tracked neuron's PETH on different days. For this analysis, we use the time window -100ms to +500ms around reach onset, which captures most reaching timepoints while excluding the majority of non-reaching timepoints. For each animal we examined the distribution of these correlation values as a function of the interval between days and compared it to a control distribution of correlations between PETHs for non-same cells (Fig 4B; see Methods for details of control distribution). We observed that the distribution of correlation values for each interval was clustered near the maximum correlation value of 1 and was far from the median of the control distribution, which was close to 0 (example animal in Fig 4B; all animals in Fig 4C). In other words, neurons are more similar to themselves over days than to other neurons. This indicates that the time course of the average responses during and around reaching remained stable in this population across an interval of five days. These findings are consistent with observations for other behaviors (Katlowitz et al., 2018; Jensen et al., 2022), and suggests that encoding for dexterous movements in rodent motor cortex is also stable.

Trial-by-trial variance in forelimb, digit, and head movements enables accurate modeling of single-trial neuronal activity in CFA M1. Paw and head movements during reaching exhibit variability across multiple kinematic dimensions such as speed, duration, the timing of wrist pronation and of the digits opening and closing. as described in Fig 1 & 2. While PETHs quantify a neuron's response for the average movement, they do not reveal whether a neuron encodes the significant trial-by-trial variance in this behavior. To assess the neural encoding of the moment-to-moment details of reach-to-grasp kinematics, we constructed families of linear models, referred to as neural encoding models, to predict the activity of individual neurons based on the extensive set of kinematic variables we extracted using keypoint tracking (schematized in Fig 5A).

It is important to note that it is possible for the mean movement to be encoded, while the trial-by-trial variance around the mean is not. This would suggest that there is a general "reach" signal in motor cortex (indicating whether the animal is performing this behavior or not), without encoding the moment-to-moment details of the constituent movements during that behavior. While individual

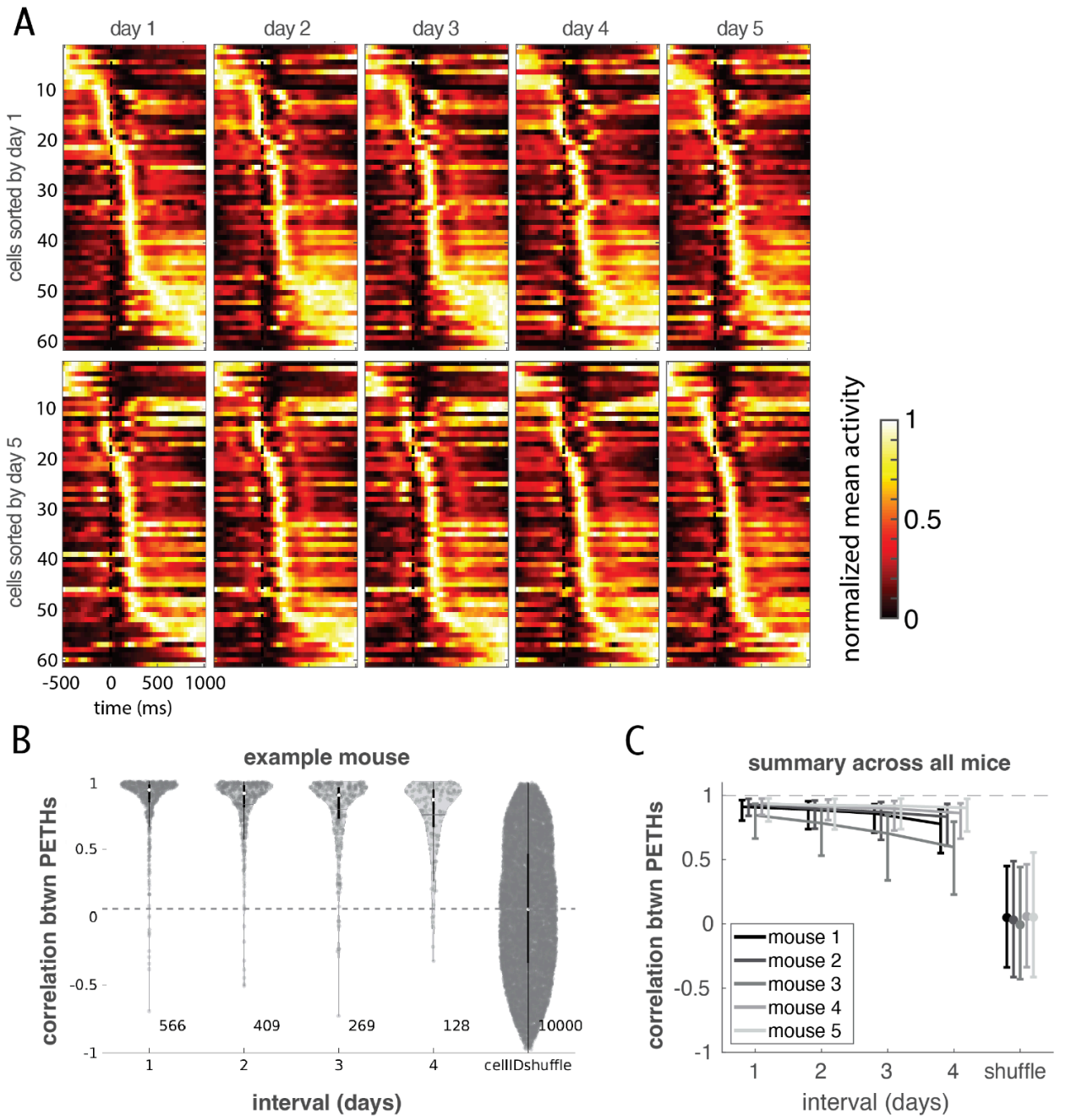

Figure 4: Reach-locked average neuronal activity patterns (PETHs) are stable across days

A) Normalized mean activity (PETHs) of neurons registered across all five days for an example mouse (mouse 4). Cells are sorted according to the time of their peak activity on either day 1 (top row) or day 5 (bottom row). Each cell's activity has been normalized separately within this window: minimum value = 0, maximum value = 1.

B) Correlation between PETHs for cells found on more than one day, as a function of the interval between the days, for an example mouse (mouse 4). Each datapoint is an observation of the same cell on two different days. Numbers to the bottom right of the violins indicate the number of observations included in that violin. For each violin, the white dot indicates the median of the distribution and the black box indicates interquartile range. Cell ID shuffle indicates the distribution of correlations between PETHs of two cells randomly chosen from any day (excluding any cell pairs that are registered as the same cell; see Methods). Gray dashed line indicates the median of this shuffle.

C) Correlation between PETHs for all mice. Datapoint is median, error bars are interquartile range across cells for that interval for that mouse.

neurons exhibit significant trial-by-trial variance around their average responses (Fig 3D), it is possible that that variance does not co-vary with behavior. If movement details are indeed encoded in motor cortex within a given session, it is also possible that the nature of that encoding is not the same across days, indicating instability (drift).

We first applied a covariate balancing subsampling procedure to the predictor sets for all five days, individually for each animal (see Methods and Supplementary Fig. 3). This step ensured that we compared the encoding of similar movements, submovements and their corresponding variability across days.

The majority of neurons on any given single day have a substantial amount of variance explained by the set of kinematic predictors compared to a timepoint-shuffled control. Specifically, 4883 out of 6105 *cell-days* (80.0%) showed significant variance explained, where each cell-day represents a single day's observation for a given neuron. This analysis included data from all days and all mice ($p = 0.05$, with Bonferroni correction for the number of cells in each dataset/field of view for that day; see Methods for details of the shuffle comparison).

This result alone should not be interpreted as definitive evidence of trial-by-trial variance encoding. All linear models in this study are timepoint-by-timepoint models, which treat each timepoint as an independent observation and regress across all observations in a session. This approach is standard for quantifying the neural encoding of movements, but it conflates encoding for the mean movement with encoding for the variability around the mean. We have already observed that individual cells exhibit time-varying average responses (PETHs in Fig 4A), which can lead to significant predictions in timepoint-by-timepoint models even in the absence of coding for variance around the mean. In other words, the time-varying average movement could explain neural activity above and beyond the timepoint shuffle control. To test for this possibility, we built another model for each cell, in which we predict the cell's single trial activity from trial-averaged movements (schematized in Fig 5B). In the predictor matrix of this trial-average model, we replace the timepoints of each reach with the reach-locked average, for each kinematic predictor separately. Approximately half of all recorded cells were significantly predicted by the trial average model (3097 out of 6105 cell-days, or 50.7%, where each day for a given cell is a separate observation (cell-day), using cells from all days for all mice; using the same shuffle comparison as for the single trial full model, see Methods for details), corresponding to 50.5% of each dataset (50.5% +/- 13.0% [17.4% 71.5%]; mean +/- std [range]). Most neurons with significant predictions in the trial-average model were also significantly predicted by the single-trial model (2972 out of 3097 cell-days, or 96.0%). Very few neurons were significantly predicted by only the trial-average model, indicating they encode only a general reach signal (125 out of 6105 cell-days, or 2.0%). In contrast, many neurons showed significant predictions in the single-trial model but not in the trial-average model, suggesting that they encode moment-to-moment details but not a general reach signal (1911 out of 6105 cell-days, or 31.3%).

Moreover, of the neurons significantly predicted by the single trial model, we found that the vast majority were better modeled by the single trial model than by the trial average model (Fig 5C-E; 4522 out of 4883 (92.6%) cell-days with significant single trial full model predictions, from all days for all mice). This finding suggests that these neurons encode single trial variance around the average reach. These results indicate that mouse motor cortex encodes not only the general "reaching" signal observed in the PETHs and trial average predictions but also the moment-to-moment details of movement. For the remainder of our analysis, we focus on the subset of neurons with significant single-trial model predictions. On average, this subset comprises 79.4% of neurons in each dataset (79.4% +/- 16.3%, [46.8% 94.2%] (mean +/- sd, [range] across datasets, n=25 datasets), see Supp Fig. 4B,C for proportion and counts by mouse and dataset).

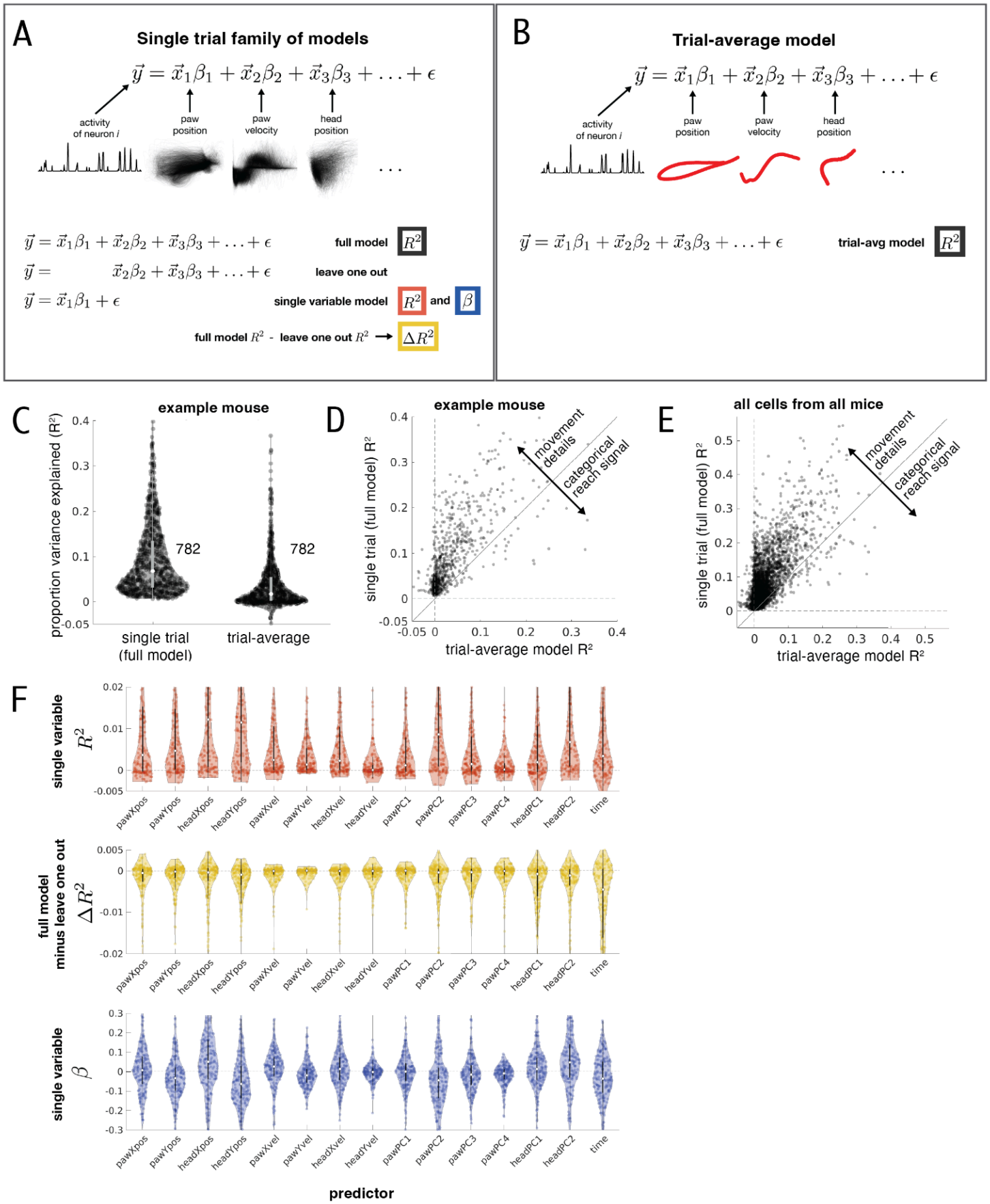
A
Single trial family of models
$\vec{y} = \vec{x}_1\beta_1 + \vec{x}_2\beta_2 + \vec{x}_3\beta_3 + \ldots + \epsilon$
activity of neuron i
paw position
paw velocity
head position
$\vec{y} = \vec{x}_1\beta_1 + \vec{x}_2\beta_2 + \vec{x}_3\beta_3 + \ldots + \epsilon$ full model R^2
$\vec{y} = \vec{x}_2\beta_2 + \vec{x}_3\beta_3 + \ldots + \epsilon$ leave one out
$\vec{y} = \vec{x}_1\beta_1 + \epsilon$ single variable model R^2 and β
full model R^2 - leave one out R^2 → ΔR^2
B
Trial-average model
$\vec{y} = \vec{x}_1\beta_1 + \vec{x}_2\beta_2 + \vec{x}_3\beta_3 + \ldots + \epsilon$
activity of neuron i
paw position
paw velocity
head position
$\vec{y} = \vec{x}_1\beta_1 + \vec{x}_2\beta_2 + \vec{x}_3\beta_3 + \ldots + \epsilon$ trial-avg model R^2
C
example mouse
proportion variance explained (R²)
782
782
single trial (full model)
trial-average
D
example mouse
single trial (full model) R²
trial-average model R²
movement details
categorical reach signal
E
all cells from all mice
single trial (full model) R²
trial-average model R²
movement details
categorical reach signal
F
single variable R^2
full model minus leave one out ΔR^2
single variable β
pawXpos
pawYpos
headXpos
headYpos
pawXvel
pawYvel
headXvel
headYvel
pawPC1
pawPC2
pawPC3
pawPC4
headPC1
headPC2
time
predictor

Figure 5: Trial-by-trial variance in forelimb, digit, and head movements allows for accurate modeling of single-trial CFA M1 neuronal activity

A) Schematic of trial-to-trial linear model families built to predict the activity of individual cells from kinematic features during reaching. The full trial-to-trial model contains all kinematic features and their observed values across trials. For each predictor, we build a leave-one-out model using all predictors except the one in question, and a single variable model, using only that predictor. The change in variance explained from the full model to the leave one out model indicates unique variance explained by that predictor.

B) Schematic of the trial-average linear model, where each trial of the kinematics is replaced by the average timecourse for that kinematic predictor, averaged across reaches, locked to reach onset (see Methods).

C) Proportion of variance explained (R-squared) by the full single trial model or the trial average model for all cells with a significant full model prediction, for an example mouse (mouse 5). Numbers to the right of each violin indicate the number of observations included in the distribution. Each datapoint corresponds to a cell on one day. Cells registered across days have multiple dots if they have significant full trial-to-trial model predictions on multiple days. For each violin, the white dot indicates the median of the distribution and the black box indicates interquartile range.

D) Same data as in C, but with the full model and trial average model R squared values plotted against each other for each cell.

E) Same as D, but for all cells with significant full model predictions for all days for all mice.

F) Top: coefficients from single variable model fits for all cells in one example session. Each datapoint corresponds to a cell; each cell has one datapoint in each violin. Middle: same as top but for the variance explained (R-squared) for single variable model fits. Bottom: same as top but for the change in variance explained (delta R-squared) for full models to leave-one-out models for each predictor.

The R squared value from the single trial model represents the maximum amount of variance that the entire predictor set can explain for each individual neuron. Next we consider the contribution of individual predictors. Due to the nature of the task and the physics of the animals' bodies, many of our kinematic variables exhibit non-zero correlations (Supplementary Fig. 4A). Consequently, the coefficients in our single trial linear model are not unique and should not be interpreted as the maximum contribution of each individual predictor (Stevenson et al., 2018). A common approach to allow direct comparison of the coefficients is to orthogonalize the predictors to remove any correlations between them. However, this approach biases the analysis towards predictors that appear earlier in the design matrix, potentially underestimating the explanatory power of predictors that come later. To avoid this, we computed two metrics to quantify the contribution of individual kinematic predictors to explain a neuron's variance (schematized in Fig. 5A): the variance explained by each predictor separately (single variable R-squared values) and its unique contribution (delta-R-squared values). 1) Single Variable Model: For each predictor, we constructed a single-variable model using only the predictor in question. The R-squared value from the single-variable model reflects the maximum variance that that predictor can contribute to the full model. 2) Leave-One-Out Model: Starting with the full model, which predicts a neuron's activity using all kinematic predictors, we created a leave-one-out model by excluding one predictor at a time and refitting the model. The drop in R-squared when a predictor is removed, referred to as the delta-R-squared, indicates the unique variance that the predictor explains for that neuron. Together, these metrics provide a comprehensive view of both the unique and total explanatory power of individual predictors in modeling neuronal activity. Additionally, the coefficients from the single variable models can serve as tuning curves, revealing how the neuron's activity is modulated in relation to variations in the kinematic predictor.

Many of the predictors exhibit nonzero single variable R squared values and delta R squared values (Fig 5F top, middle), indicating that these movement variables are encoded by single cells in motor cortex. Additionally, distributions of single variable model coefficients (betas) are dispersed away from zero (Fig 5F bottom). Notably, these observations apply not only to paw-related predictors but also to head-related predictors. Head and paw variables are correlated (Supplementary Fig. 4A), which can account for some of the single variable R-squared values. However,

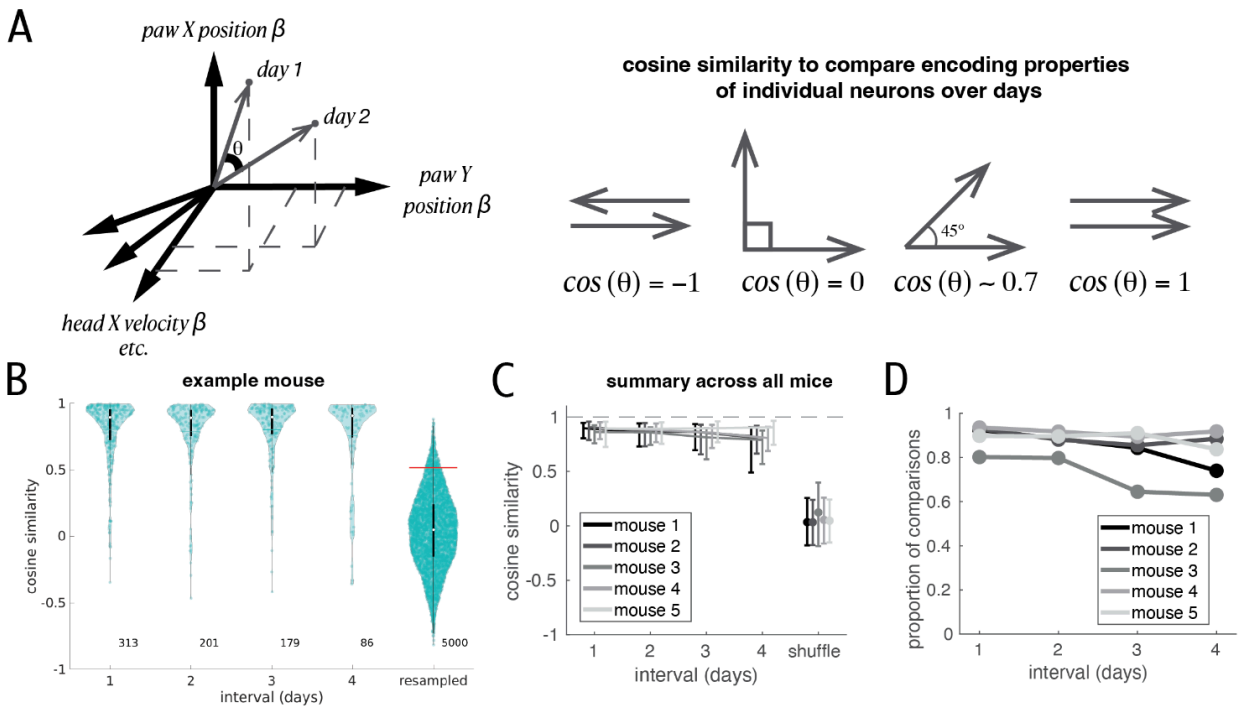

Figure 6: Linear model encoding profiles of single cells are stable across days

A) Schematic of method: using cosine similarity to compare vectors of encoding coefficients (betas) over days for individual cells

B) Cosine similarities of coefficient vectors as a function of the interval between observations for an example mouse (mouse 5). Each datapoint corresponds to a comparison of observations of a single cell on two different days. Numbers to the bottom right of the violins indicate the number of observations for that interval. For each violin, the white dot indicates the median of the distribution and the black box indicates interquartile range. Resampled shuffle control built by generating pseudo-vectors where each coefficient is drawn from the pool of coefficients observed across all neurons for that predictor within that day. Coefficients within the same pseudo-vector are drawn independently. The resampled shuffle distribution contains the same number of comparisons for each of the 10 possible pairs of the 5 days of study (5000 total, 500 per day pair). Red line indicates the 95th percentile cutoff of this distribution.

C) Same as B, but for all animals. Data is median, error bars are interquartile range across cells for that interval for that mouse.

D) Proportions of coefficient vector comparisons for a given interval that are statistically significant using the resampled shuffle distribution as a control.

several head variables have non-zero delta R-squared values for many neurons, indicating that the position and orientation of the head are encoded in forelimb motor cortical activity in addition to paw-related predictors. These findings are further evidence that mouse motor cortex encodes both a general "reach" signal (observed in the PETHs and the trial average model predictions) and the moment-to-moment details of the reaching movements. These details include not only the position of the paw but its shape and orientation (posture of the digits) and the position and orientation of the head.

Linear model encoding profiles of single cells are stable across days. To assess whether the encoding of trial-by-trial variability is consistent across days, we compared components of the model fits across each day on which a neuron was recorded. A common approach to quantifying the stability of activity-behavior relationships is to assess how well a model trained on one day performs when applied to data from other days, often as a function of the number of days between the training and test sessions (e.g., Driscoll et al., 2017). When predictors are correlated, as in our study, there is an infinite set of possible coefficient combinations that can produce the same optimal fit (Stevenson et al., 2018). If the

correlations between predictors, or between predictors and omitted (but relevant) variables, change even slightly from the training day to the test day, the coefficients derived from the training day may no longer accurately reflect the predictor relationships on the test day. This occurs even if the underlying relationships of interest—such as the link between individual predictors and the predicted variable (e.g., a neuron's fluorescence)—remain stable. This would result in a drop in performance of the model on the test day. If predictor correlations are not carefully accounted for, applying fitted models across days may mistakenly suggest that activity-behavior relationships for individual predictors are changing, when in fact, it is the relationships among the predictors that are shifting.

To compare the contribution of the predictors across days, we instead compared the collection of single variable model coefficients (betas), referred to as a neuron's encoding profile. Specifically, for a given cell, we concatenated the coefficients for all predictors from one day into a vector and calculated the cosine similarity between a neuron's vectors from different days (Fig 6A, left). Cosine similarity is simply the cosine of the angle between the two vectors and ranges from -1 to 1, indicating antiparallel and parallel vectors, respectively (Fig. 6A, right). This approach compares not the exact value of each coefficient over days, but instead the relationships between the coefficients for all predictors. Thus it provides a holistic comparison of a neuron's tuning on different days. Importantly, we are comparing the coefficients of the single variable models, rather than the coefficients of the full model, for all the reasons described in the previous paragraph. We only consider the stability of encoding of cells with significant full single trial model predictions; for cells with significant predictions on some days but not all days, we only compare encodings from days with significant predictions. We found that across intervals of up to four days apart, cosine similarity values are high for the vast majority of comparisons of same-cell encoding profiles (example mouse in Fig. 6B, all mice in Fig. 6C). Of the 4926 comparisons across all intervals for all mice, 4354 were significant (88.4%) relative to their respective resampled control distribution (Fig. 6D for breakdown by mouse and interval; see Methods for details of control distribution). We also performed the same comparison with encoding profiles made up of the collection of R-squared values (single-variable model R-squared values, full model R-squared, and delta R-squareds) and again observe significance for the majority of comparisons (3222 out of 4926 (65.4%); see Supplementary Fig. 5). These findings demonstrate that the encoding of moment-to-moment movement details in motor cortical activity remains largely stable at the single-neuron level across multiple days.

Discussion

Here, we investigated how the motor cortex encodes moment-to-moment movements of the paw, digits, and head during a self-initiated, skilled reach-to-grasp task, and whether this encoding changes over days. Consistent with prior studies on the stability of motor systems (Jensen et al., 2022; Katlowitz et al., 2018; Liberti et al., 2016), our findings demonstrate that average responses (PETHs) in M1 remain highly stable over time. We extend previous work by recording motor cortex activity during the performance of a challenging motor skill characterized by many repetitions within a day and high reach-to-reach variability, features that make the skill more reflective of real-world motor behaviors. We find that neurons in the mouse motor cortex encode both general reaching signals and the detailed, moment-to-moment movements of the head, forelimb, and digits during reach-to-grasp behaviors. This encoding remains remarkably stable over days, providing a reliable neural foundation for producing skilled movements consistently over time.

Quantifying neural encoding during naturalistic behaviors with high trial-to-trial variability poses significant challenges which we overcame through a series of carefully designed approaches. First, we used linear encoding models that assume a consistent linear relationship between movement variables and neural activity within a day,

rather than relying on assumptions of stereotyped movements across trials. This analytical flexibility allowed us to account for the inherent variability of naturalistic behaviors. Second, we designed the task to balance variability with sufficient trial counts. Key elements of the task, such as requiring animals to reposition themselves after each bout of reaches, randomizing seed positions across multiple locations, and omitting penalties for failed attempts, encouraged trial-to-trial variability while preventing the development of stereotyped automatized movement patterns. The task's inherent difficulty likely further contributed to the diversity in movements. Third, we captured these movements at a fine-grained resolution using high-speed, high-resolution cameras, allowing us to precisely quantify trial-to-trial variability above the noise threshold. This detailed behavioral data was crucial for accurately identifying variations in movement. Finally, to ensure that observed changes in encoding were not confounded by behavioral shifts over days, we implemented a covariate balancing procedure before fitting linear models. This approach allowed us to quantify encoding within consistent regions of the high-dimensional kinematic parameter sets across days, effectively isolating neural changes from behavioral variability. Together, these strategies enabled us to rigorously estimate and compare neural encoding across days in a statistically robust manner. Our findings not only highlight the stability of motor cortex encoding during complex, variable behaviors but also underscore the value of incorporating detailed behavioral analyses into neural encoding studies. We hypothesize that similarly rigorous approaches to quantifying and incorporating behavior into encoding models, including movements not explicitly instructed by the task (Sadeh and Clopath, 2022; Musall, Kaufman, et al., 2019; Stringer et al., 2019), could reveal stable encoding throughout motor circuitry and potentially in sensory and association brain areas as well.

The peri-event time histogram (PETH) captures the time-varying average response of a neuron across repeated instances of a specific behavioral or task-related event. In this study, the PETH serves as a categorical descriptor of "reaching," since we aligned neural responses to the onset of each reach. Our analysis reveals a consistent spatiotemporal pattern of neural activation (PETH) during reaching that was conserved across days. For neurons reliably tracked over time, the vast majority exhibited stable average responses, consistent with findings in rats (Jensen et al., 2022) and zebra finches (Katlowitz et al., 2018; Liberti et al., 2016). The observed stability in the PETH for a task requiring dexterous digit control in mice under naturalistic conditions reinforces the idea that the rodent motor cortex can sustain stable encoding of learned movements, irrespective of task complexity. Collectively these findings strongly suggest that stable behaviors arise from the consistent contributions of individual neurons within the motor circuit.

While the PETH provides an estimate of a neuron's mean activity during reaching, it does not account for trial-to-trial variability in the behavior, which could offer additional insights into the dynamics of neural encoding. We used single-trial encoding models to investigate whether the significant trial-to-trial variability in movements is encoded in the motor cortex and whether this encoding remains stable over days. These models linked the activity of individual neurons to a carefully curated set of 14 kinematic variables, capturing the movements of specific body parts (e.g., position and velocity of the paw and head, as well as the conformation of the digits and the orientation of the paw and head). While more complex models could likely explain additional variance, they often come at the cost of interpretability. Despite using straightforward linear approaches, we found that moment-to-moment movement details during reach-to-grasp explained a statistically significant fraction of the variance in most recorded neurons. Furthermore, these relationships remained stable across multiple days, representing the longest duration over which such detailed single-cell encoding of moment-to-moment movement has been examined.

The finding that the motor cortex encodes single-trial variability in reach-to-grasp movements suggests a role in the online monitoring and control of this sensory-driven behavior, operating as part of broader sensorimotor loops involving other brain regions (Sauerbrei et al., 2020; Wagner et al., 2019). The linear models showed that the principal components of paw markers, which reflect wrist and digit states, contributed explanatory power comparable to paw position, which reflects the state of proximal joints such as the shoulder and elbow. Although one causal manipulation study suggests that the rostral forelimb area (RFA) is more involved in grasping than the caudal forelimb area (CFA) (Wang et al., 2017), our finding of digit encoding supports CFA involvement. Further investigation is needed to clarify how digit-related signals in CFA and RFA might work together to coordinate grasping.

Interestingly, the position and orientation of the head also contributed explanatory power comparable to paw- and digit-related predictors in the encoding model. This could reflect two potential (and not mutually exclusive) explanations: first, head-related variables may be correlated with omitted kinematic variables that are more directly linked to neural activity (Stevenson et al., 2018). Alternatively, head position and orientation may themselves be directly relevant to computations in the motor cortex, providing crucial information about body positioning relative to the reach port and pedestal. This latter explanation implies a role for egocentric planning signals in guiding the paw to the pedestal, as observed in encoding of whole body posture in rats (Mimica et al., 2018). These findings highlight the importance of examining how multiple body parts are encoded across various brain areas, rather than limiting the focus to regions traditionally associated with specific movements. A comprehensive understanding of whole-body movement generation requires this broader perspective. Furthermore, to uncover how the brain produces the flexible, context-dependent, and coordinated movements essential for daily human activities, it is crucial to study comparable behaviors in animal models. The reach-to-grasp behavior analyzed here exemplifies such complexity, combining whole-body posture with precise forelimb and digit control. This task closely parallels primate reaching and grasping behaviors (Sacrey et al., 2009; Whishaw et al., 1992; Guo et al., 2015; Becker et al., 2020; Whishaw et al., 2017) and serves as a powerful model for investigating the neural mechanisms underlying these intricate actions.

References

An, X., Matho, K., Li, Y., Mohan, H., Xu, X. H., Whishaw, I. Q., Kepecs, A., & Huang, Z. J. (2024). A cortical circuit for orchestrating oromanual food manipulation. bioRxiv. https://doi.org/10.1101/2022.12.03.518964

Andermann, M. L., Gilfoy, N. B., Goldey, G. J., Sachdev, R. N. S., Wölfel, M., McCormick, D. A., Reid, R. C., & Levene, M. J. (2013). Chronic Cellular Imaging of Entire Cortical Columns in Awake Mice Using Microprisms. *Neuron*, *80*(4), 900–913. https://doi.org/10.1016/j.neuron.2013.07.052

Barrett, J. M., Raineri Tapies, M. G., & Shepherd, G. M. G. (2020). Manual dexterity of mice during food-handling involves the thumb and a set of fast basic movements. *PLOS ONE*, *15*(1), e0226774. https://doi.org/10.1371/journal.pone.0226774

Bauer, J., & Rose, T. (2021). Mouse vision: Variability and stability across the visual processing hierarchy. *Current Biology*, *31*(19), R1129–R1132. https://doi.org/10.1016/j.cub.2021.08.071

Becker, M. I., Calame, D. J., Wrobel, J., & Person, A. L. (2020). Online control of reach accuracy in mice. *Journal of Neurophysiology*, *124*(6), 1637–1655. https://doi.org/10.1152/jn.00324.2020

Billard, A., & Kragic, D. (2019). Trends and challenges in robot manipulation. *Science*, *364*(6446), eaat8414. https://doi.org/10.1126/science.aat8414

Chambers, A. R., & Rumpel, S. (2017). A stable brain from unstable components: Emerging concepts and implications for neural computation. *Neuroscience*, *357*, 172–184. https://doi.org/10.1016/j.neuroscience.2017.06.005

Chen, C.-C., Gilmore, A., & Zuo, Y. (2014). Study Motor Skill Learning by Single-pellet Reaching Tasks in Mice. *Journal of Visualized Experiments : JoVE*, *85*. https://doi.org/10.3791/51238

Chestek, C. A., Batista, A. P., Santhanam, G., Yu, B. M., Afshar, A., Cunningham, J. P., Gilja, V., Ryu, S. I., Churchland, M. M., & Shenoy, K. V. (2007). Single-Neuron Stability during Repeated Reaching in Macaque Premotor Cortex. *Journal of*

Neuroscience, *27*(40), 10742–10750. https://doi.org/10.1523/JNEUROSCI.0959-07.2007

Clopath, C., Bonhoeffer, T., Hübener, M., & Rose, T. (2017). Variance and invariance of neuronal long-term representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1715), 20160161. https://doi.org/10.1098/rstb.2016.0161

Dhawale, A. K., Poddar, R., Wolff, S. B., Normand, V. A., Kopelowitz, E., & Ölveczky, B. P. (2017). Automated long-term recording and analysis of neural activity in behaving animals. *ELife*, *6*, e27702. https://doi.org/10.7554/eLife.27702

Disse, G. D., Nandakumar, B., Pauzin, F. P., Blumenthal, G. H., Kong, Z., Ditterich, J., & Moxon, K. A. (2023). Neural ensemble dynamics in trunk and hindlimb sensorimotor cortex encode for the control of postural stability. *Cell Reports*, *42*(4), 112347. https://doi.org/10.1016/j.celrep.2023.112347

Dombeck, D. A., Graziano, M. S., & Tank, D. W. (2009). Functional Clustering of Neurons in Motor Cortex Determined by Cellular Resolution Imaging in Awake Behaving Mice. *Journal of Neuroscience*, *29*(44), 13751–13760. https://doi.org/10.1523/JNEUROSCI.2985-09.2009

Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N., & Harvey, C. D. (2017). Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. *Cell*, *170*(5), 986-999.e16. https://doi.org/10.1016/j.cell.2017.07.021

Driscoll, L. N., Duncker, L., & Harvey, C. D. (2022). Representational drift: Emerging theories for continual learning and experimental future directions. *Current Opinion in Neurobiology*, *76*, 102609. https://doi.org/10.1016/j.conb.2022.102609

Flint, R. D., Scheid, M. R., Wright, Z. A., Solla, S. A., & Slutzky, M. W. (2016). Long-Term Stability of Motor Cortical Activity: Implications for Brain Machine Interfaces and Optimal Feedback Control. *Journal of Neuroscience*, *36*(12), 3623–3632. https://doi.org/10.1523/JNEUROSCI.2339-15.2016

Fraser, G. W., & Schwartz, A. B. (2012). Recording from the same neurons chronically in motor cortex. *Journal of Neurophysiology*, *107*(7), 1970–1978. https://doi.org/10.1152/jn.01012.2010

Galiñanes, G. L., Bonardi, C., & Huber, D. (2018). Directional Reaching for Water as a Cortex-Dependent Behavioral Framework for Mice. *Cell Reports*, *22*(10), 2767–2783. https://doi.org/10.1016/j.celrep.2018.02.042

Gallego, J. A., Perich, M. G., Chowdhury, R. H., Solla, S. A., & Miller, L. E. (2020). Long-term stability of cortical population dynamics underlying consistent behavior. *Nature Neuroscience*. https://doi.org/10.1038/s41593-019-0555-4

Ganguly, K., & Carmena, J. M. (2009). Emergence of a Stable Cortical Map for Neuroprosthetic Control. *PLoS Biology*, *7*(7), e1000153. https://doi.org/10.1371/journal.pbio.1000153

Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., Taxidis, J., Najafi, F., Gauthier, J. L., Zhou, P., Khakh, B. S., Tank, D. W., Chklovskii, D. B., & Pnevmatikakis, E. A. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. *ELife*, *8*, e38173. https://doi.org/10.7554/eLife.38173

Guo, J.-Z., Graves, A. R., Guo, W. W., Zheng, J., Lee, A., Rodríguez-González, J., Li, N., Macklin, J. J., Phillips, J. W., Mensh, B. D., Branson, K., & Hantman, A. W. (2015). Cortex commands the performance of skilled movement. *ELife*, *4*. https://doi.org/10.7554/eLife.10774

Huber, D., Gutnisky, D. A., Peron, S., O'Connor, D. H., Wiegert, J. S., Tian, L., Oertner, T. G., Looger, L. L., & Svoboda, K. (2012). Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature*, *484*(7395), 473–478. https://doi.org/10.1038/nature11039

Jensen, K. T., Kadmon Harpaz, N., Dhawale, A. K., Wolff, S. B. E., & Ölveczky, B. P. (2022). Long-term stability of single neuron activity in the motor system. *Nature Neuroscience*, *25*(12), 1664–1674. https://doi.org/10.1038/s41593-022-01194-3

Katlowitz, K. A., Picardo, M. A., & Long, M. A. (2018). Stable Sequential Activity Underlying the Maintenance of a Precisely Executed Skilled Behavior. *Neuron*, *98*(6), 1133-1140.e3. https://doi.org/10.1016/j.neuron.2018.05.017

Koh, N., Ma, Z., Sarup, A., Kristl, A. C., Agrios, M., Young, M., & Miri, A. (2024). Selective direct motor cortical influence during naturalistic climbing. bioRxiv. https://doi.org/10.1101/2023.06.18.545509

Liberti, W. A., Markowitz, J. E., Perkins, L. N., Liberti, D. C., Leman, D. P., Guitchounts, G., Velho, T., Kotton, D. N., Lois, C., & Gardner, T. J. (2016). Unstable neurons underlie a stable learned behavior. *Nature Neuroscience*, *19*(12), 1665–1671. https://doi.org/10.1038/nn.4405

Liberti, W. A., Schmid, T. A., Forli, A., Snyder, M., & Yartsev, M. M. (2022). A stable hippocampal code in freely flying bats. *Nature*, *604*(7904), 98–103. https://doi.org/10.1038/s41586-022-04560-0

Lütcke, H., Margolis, D. J., & Helmchen, F. (2013). Steady or changing? Long-term monitoring of neuronal population activity. *Trends in Neurosciences*, *36*(7), 375–384. https://doi.org/10.1016/j.tins.2013.03.008

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

Micou, C., & O'Leary, T. (2023). Representational drift as a window into neural and behavioural plasticity. *Current Opinion in Neurobiology*, *81*, 102746. https://doi.org/10.1016/j.conb.2023.102746

Mimica, B., Dunn, B. A., Tombaz, T., Bojja, V. P. T. N. C. S., & Whitlock, J. R. (2018). Efficient cortical coding of 3D posture in freely behaving rats. *Science*, *362*(6414), 584–589. https://doi.org/10.1126/science.aau2013

Miri, A., Warriner, C. L., Seely, J. S., Elsayed, G. F., Cunningham, J. P., Churchland, M. M., & Jessell, T. M. (2017). Behaviorally Selective Engagement of Short-Latency Effector Pathways by Motor Cortex. *Neuron*, *95*(3), 683-696.e11. https://doi.org/10.1016/j.neuron.2017.06.042

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677–1686. https://doi.org/10.1038/s41593-019-0502-4

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless

pose estimation across species and behaviors. *Nature Protocols*, *14*(7), 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Omlor, W., Sipilä, P., Wahl, A.-S., Lütcke, H., Laurenczy, B., Chen, I.-W., Sumanovski, L. T., 't Hoff, M. van, Bethge, P., Voigt, F. F., Schwab, M. E., & Helmchen, F. (2019). Context-dependent limb movement encoding in neuronal populations of motor cortex. bioRxiv. https://doi.org/10.1101/588129

Peters, A. J., Chen, S. X., & Komiyama, T. (2014). Emergence of reproducible spatiotemporal activity during motor learning. *Nature*, *510*(7504), 263–267. https://doi.org/10.1038/nature13235

Peters, A. J., Lee, J., Hedrick, N. G., O'Neil, K., & Komiyama, T. (2017). Reorganization of corticospinal output during motor learning. *Nature Neuroscience*, *20*(8), 1133–1141. https://doi.org/10.1038/nn.4596

Resendez, S. L., Jennings, J. H., Ung, R. L., Namboodiri, V. M. K., Zhou, Z. C., Otis, J. M., Nomura, H., McHenry, J. A., Kosyk, O., & Stuber, G. D. (2016). Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. *Nature Protocols*, *11*(3), 566–597. https://doi.org/10.1038/nprot.2016.021

Rokni, U., Richardson, A. G., Bizzi, E., & Seung, H. S. (2007). Motor Learning with Unstable Neural Representations. *Neuron*, *54*(4), 653–666. https://doi.org/10.1016/j.neuron.2007.04.030

Sacrey, L.-A. R., Alaverdashvili, M., & Whishaw, I. Q. (2009). Similar hand shaping in reaching-for-food (skilled reaching) in rats and humans provides evidence of homology in release, collection, and manipulation movements. *Behavioural Brain Research*, *204*(1), 153–161. https://doi.org/10.1016/j.bbr.2009.05.035

Sadeh, S., & Clopath, C. (2022). Contribution of behavioural variability to representational drift. *ELife*, *11*, e77907. https://doi.org/10.7554/eLife.77907

Sauerbrei, B. A., Guo, J.-Z., Cohen, J. D., Mischiati, M., Guo, W., Kabra, M., Verma, N., Mensh, B., Branson, K., & Hantman, A. W. (2020). Cortical pattern generation during dexterous movement is input-driven. *Nature*, *577*(7790), 386–391. https://doi.org/10.1038/s41586-019-1869-9

Sheintuch, L., Rubin, A., Brande-Eilat, N., Geva, N., Sadeh, N., Pinchasof, O., & Ziv, Y. (2017). Tracking the Same Neurons across Multiple Days in Ca2+ Imaging Data. *Cell Reports*, *21*(4), 1102–1115. https://doi.org/10.1016/j.celrep.2017.10.013

Stevenson, I. H. (2018). Omitted Variable Bias in GLMs of Neural Spiking Activity. *Neural Computation*, *30*(12), 3227–3258. https://doi.org/10.1162/neco_a_01138

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, *364*(6437), eaav7893. https://doi.org/10.1126/science.aav7893

Wagner, M. J., Kim, T. H., Kadmon, J., Nguyen, N. D., Ganguli, S., Schnitzer, M. J., & Luo, L. (2019). Shared Cortex-Cerebellum Dynamics in the Execution and Learning of a Motor Task. *Cell*, *177*(3), 669-682.e24. https://doi.org/10.1016/j.cell.2019.02.019

Wang, X., Liu, Y., Li, X., Zhang, Z., Yang, H., Zhang, Y., Williams, P. R., Alwahab, N. S. A., Kapur, K., Yu, B., Zhang, Y., Chen, M., Ding, H., Gerfen, C. R., Wang, K. H., & He, Z. (2017). Deconstruction of Corticospinal Circuits for Goal-Directed Motor Skills. *Cell*, *171*(2), 440-455.e14. https://doi.org/10.1016/j.cell.2017.08.014

Whishaw, I. Q., Faraji, J., Kuntz, J., Mirza Agha, B., Patel, M., Metz, G. A. S., & Mohajerani, M. H. (2017). Organization of the reach and grasp in head-fixed vs freely-moving mice provides support for multiple motor channel theory of neocortical organization. *Experimental Brain Research*, *235*(6), 1919–1932. https://doi.org/10.1007/s00221-017-4925-4

Whishaw, I. Q., Pellis, S. M., & Gorny, B. P. (1992). Skilled reaching in rats and humans: evidence for parallel development or homology. *Behavioural Brain Research*, *47*(1), 59–70. https://doi.org/10.1016/S0166-4328(05)80252-9

Zhou, P., Resendez, S. L., Rodriguez-Romaguera, J., Jimenez, J. C., Neufeld, S. Q., Giovannucci, A., Friedrich, J., Pnevmatikakis, E. A., Stuber, G. D., Hen, R., Kheirbek, M. A., Sabatini, B. L., Kass, R. E., & Paninski, L. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *ELife*, *7*, e28728. https://doi.org/10.7554/eLife.28728

Zhu, F., Grier, H. A., Tandon, R., Cai, C., Agarwal, A., Giovannucci, A., Kaufman, M. T., & Pandarinath, C. (2022). A deep learning framework for inference of single-trial neural population dynamics from calcium imaging with subframe temporal resolution. *Nature Neuroscience*, *25*(12), 1724–1734. https://doi.org/10.1038/s41593-022-01189-0

Acknowledgements

The authors would like to thank Gabriella Wheeler Fox, Zach Morrissey, Siwei Wang, Harold Rockwell, Sunnie Hong, Matthew Kaufman and David McLean for their feedback on the manuscript. The authors would also like to thank Chadbourne M. B. Smith, Steven Redford, and Gabriella Wheeler Fox for technical assistance. This work was supported by National Science Foundation Graduate Research Fellowship Program (NSF GRFP) fellowship DGE-1746045 (E.A.d.); and National Institutes of Health grants R01-NS094184 (J.N.M). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Methods

Mice

All experimental procedures were approved by the University of Chicago Institutional Animal Care and Use Committee and were performed in compliance with the Guide for the Care and Use of Laboratory Animals. Data were collected from Ai148D mice of either sex (n = 2 female, n = 3 male; Jax strain 030328). The line was originally obtained from Jackson Labs and bred in house. Mice were between 10 and 12 weeks old at the start of training.

Surgical procedures

Animals underwent three surgical procedures: AAV-Cre injections, prism implantation, and baseplate implantation. For all surgeries, anesthesia was induced with isoflurane (induction at 4%, maintenance at 1–1.5%). During AAV injection, two small burr-hole craniotomies were drilled to allow insertion of a micropipette and injection of virus. Injections were made at three depths (650, 400, and 200um below the surface) and in two locations targeting caudal forelimb area (0.25mm anterior and 1.25mm lateral to bregma, and again at 0.25mm anterior and 1.75mm lateral to bregma). After injections, the burr holes were sealed with dental cement and the skin sutured. Animals recovered for a minimum of seven days before prism implantation. During prism implantation, a 2mm diameter craniotomy was drilled, centered at 0.25mm anterior and 1.5mm lateral to bregma. A stereotaxically mounted scalpel was used to make a 1mm wide incision running along the medial lateral axis, centered at the same coordinates as the craniotomy; the incision ran to a depth of 800um. The prism was then inserted into the tissue along this incision to the same depth, with the angled part of the prism directed posteriorly, meaning the resulting field of view is anterior to the site of insertion. The remaining surface of the tissue was covered with Kwil-sil and the prism was cemented in place. Animals recovered for a minimum of seven days before baseplate implantation. During baseplate implantation, a plastic baseplate was cemented above the prism and grin lens, placed to allow the best alignment of the microscope with the focal plane of the prism. This final procedure is quick and non-invasive. Animals recovered for a minimum of two days before the start of food restriction and habituation in the arena.

Behavioral training

The entire behavioral experience consisted of three parts: habituation, shaping, and training. During all phases, the mice were briefly (~1min) anesthetized with isoflurane so that the miniscope could be attached to their headplates. Mice were then placed in the arena and allowed to recover completely (defined by showing unencumbered, normal ambulatory movement), at least 5 minutes. The arena was 15cm x 15cm with a reach port (6mm wide opening that ran from the floor up ~8cm) located in the center of one of the walls. The wall with the reach port was slanted at a 60 degree angle relative to the floor to enable the mice to maneuver better with the miniscopes on their heads.

Food restriction, Habituation and Shaping

At least one full day after headplate placement, mice began food restriction, in which they received a limited amount of pellet food, titrated to maintain their weight between 85% and 90% of their baseline weight. Mice were food restricted for at least one week before habituation began, sometimes longer in order to achieve a weight between 85% and 90% of baseline. Mice were single housed throughout food restriction, habituation, and training.

Mice were habituated to the training arena and the miniscope for at least seven days before shaping began. During habituation, mice were placed in the behavioral arena for twenty to thirty minutes, once per day. At the start of habituation, twenty millet seeds were placed on the ground inside the arena right next to the reach port, to encourage an association between the reach port and food.

During shaping, mice were again placed in the behavioral arena for twenty to thirty minutes, once per day. Instead of being presented inside the arena, seeds were presented in two places outside the arena: approximately ten seeds were presented on the ground outside the arena immediately in front of the reach port. These were reachable by the animal's tongue. Seeds were also presented in a low square trough outside the arena near the reach port. The trough was positioned a few millimeters laterally to the center of the reach port, such that the animal would only be able to successfully reach the trough with its contralateral forelimb. All mice were shaped and trained to use their left paws. The goal of shaping was to train the animal to associate reaching its forelimb outside the arena with getting food. Functionally, during shaping, the mice would scoop or knock seeds out of the trough and onto the ground, then lick them up with their tongues. Any large magnitude forelimb movement through the reach port was considered a reach. In order to move from shaping to training, mice needed to make at least 50 reaches with their left forelimb in a window of twenty minutes, and at least 80% of their reaches in that window needed to be with their left forelimb. This shaping procedure is based on that described in Chen et al., 2014. Shaping was performed daily until mice achieved these criteria. A given shaping session was not stopped when animals reached

these requirements, and often on the last shaping session before training, mice reached 75-100 times with their left forelimbs. If a mouse achieved this criteria in a shaping session, training began the next day. If, on the first training day, the animal did not show interest in reaching for the seeds on the pedestal, or reached very few times, the training session was converted to another shaping session, to reinforce the association between using their forelimbs and getting seeds. The animal was held to the same requirements as before for moving on to training.

We found that on the first day of shaping there was a critical moment requiring manual involvement. Sometimes, the mice initially reach out of the port with the correct limb but do not make contact with the seed trough or do not make strong enough contact to knock any seeds out. In the shaping phase the mice are quick to give up reaching entirely if their initial reaches are unsuccessful. We started manually knocking seeds out of the trough the first one or two times that the animal reached outside the arena with their left paw. This was generally enough.

Training

During training, seeds were only presented to the mouse on the seed delivery pedestals, which could only hold one seed each. There were two different kinds of seed delivery methods used for the mice included in this study. Mice only ever experienced one seed delivery method throughout their training.

For two of the mice, there were three permanent seed delivery pedestals positioned in an arc outside the reach port. Their positioning was offset to the mouse's right such that the pedestals were not reachable with the right forelimb. Seeds were delivered manually by the experimenter. Mice were required to turn away from the seed port and walk towards the back of the arena in order for a new seed to be delivered, thus initiating a new "trial". Only one seed was placed per trial. The order of seed placement was random but counterbalanced such that there was approximately the same number of deliveries per pedestal in a given session (matlab code).

For one of the mice, the same three permanent seed delivery pedestals were used, but seeds were only placed on one of the pedestals (the center one, equivalent to the lateral pedestal in Fig 1C).

For the final two mice, there was a single mobile seed delivery pedestal which was moved to two different positions throughout the session by a linear actuator. Their positioning was offset to the mouse's right such that the pedestals were not reachable with the right forelimb. For these mice, seeds were delivered automatically: the pedestal was lowered into a large trough of seeds and then lifted again, having caught a seed (see website for details on the automated seed delivery machinery). The order of the two delivery positions was random in a given session (python arduino code), a Bernoulli random variable. To trigger the delivery of another seed, mice had to break an infrared beam at the back of the arena.

For the manual seed delivery, if mice turned away to trigger another seed delivery, but the seed was present, the seed was left in position three times before that seed was removed and a new seed was placed. For the automated seed delivery, there was no timeout for unsuccessful reaches and no requirement that the seed be removed from the pedestal before triggering another seed delivery.

On a given training day, the mice were in the arena for between 75 and 90 minutes. They performed the task in chunks, where each chunk consisted of 3-5 minutes of seed presentations followed by 2-4 minutes of break (no seeds delivered regardless of behavior). There were 8-12 chunks per session. In total, each mouse reached (performed the behavior, was allowed to have seeds delivered) for up to 40 minutes per day. Training was performed at approximately the same time each day for a given mouse. Mice were trained for a minimum of 24 days. Imaging took place every training day. For each mouse, five sessions from late in training were used in this study, generally starting between day 20 and 22.

Behavioral video recording and kinematic reconstruction

Videography and keypoint extraction

We took high speed, color video of the mouse during reaching, at a frame rate of 200Hz, using a Flir 16S2C Blackfly camera. All kinematic data used in this paper is captured in 2D, with a camera that is at an approximately 30 degree angle from head on, offset to the mouse's left to best view the left forelimb and paw. Additionally, the camera view is angled approximately 20 degrees from above. We used DeepLabCut to extract a variety of kinematic features from this video data (Mathis et al., 2018; Nath et al., 2019). After extracting keypoints using DeepLabCut, kinematic data was further processed with custom MATLAB software. Points that were identified as poorly labeled (because of large distances traveled from the previous timepoint, or because markers that are anatomically

close together were too far apart) were linearly interpolated over. Data was then smoothed: missing frames were linearly interpolated over, we smoothed using a butterworth filter (4th order, low-pass filter with 50Hz cutoff), then removed the interpolated (originally missing) frames.

Alignment across days

The position of multiple non-moving features of the arena (referred to as static markers) were extracted from the videos and used to align kinematic data across days. We adjusted for translation only: the x and y position of a static marker was subtracted from all the kinematic variables. Generally this translation was minimal because the cameras were fixed in place relative to the arena. We also extracted the position of the pedestal and the presence of a seed for each reach automatically from the video data.

Calculating kinematic variables of interest

We calculated the centroid of the paw as the average position of seven markers (wrist, digits 1-3 bases and middles). The centroid of the head was calculated as the average of all three markers on the head (nose, front corner of the eye, and corner of the microscope). See Fig 1A and C for visualization of these markers. To calculate a low dimensional estimate of paw shape and orientation, we subtracted the paw centroid position from the seven markers on the digits and wrist, in x and y separately, then performed pca on these 14 dimensions (seven markers, in x and y). We projected data onto the first 4 dimensions, then included it in the linear model. We performed the same procedure for the three markers on the head to capture an estimate of head orientation, keeping 2 principal components of the 6 total to include in the linear model. We pooled aligned data across all five days before performing PCA, such that the PCA values were meaningful relative to each other across days. We calculate velocity from the position data using the following approach: we first calculate the difference between position in consecutive timepoints and divide that by the amount of time between timepoints (generally 5ms, but not always). We then treat that as the velocity at the halfway point between the observed timepoints and use linear interpolation to infer the velocity at the observed timepoints.

Interpolation

For the linear encoding model analysis, kinematic data (200Hz) was interpolated to the timepoints of the neural data (30Hz) using the MATLAB function interp1.

Segmenting reaches

Reaches are defined by paw position relative to the reach port: the paw was considered outside the arena when it was to the left of the corner of the right wall of the reach port (right magenta dot in Fig 1A), and inside the arena when it was to the right. The onset of a reach was identified as the moment the paw left the arena and similarly the end of the reach was identified as the timepoint the paw re-entered the arena. By this definition, reaches often consist of one out and back motion, but could include multiple out and back motions if the paw did not fully re-enter the arena in between. Reaches were clustered into multiple reaches in quick succession, followed by a longer break during which the animal left the reach port to trigger the delivery of another seed (see Fig 1F top for illustration). Bout starts were identified using the inter-reach interval: an inter-reach interval of greater than 1 second separated two bouts. For some mice in some sessions the threshold between bouts was shorter (down to 700ms), depending on the distribution of inter-reach intervals.

The three possible pedestal positions are shown in Fig 1B: medial, lateral, and far lateral. For one of the mice (mouse 3), seeds were presented in only one of the pedestal positions (lateral). For two of the mice, seeds were presented in any of the three positions, in random, counterbalanced order. For the far lateral position the animals showed no success and it was deemed too far lateral for them to reach from the port. The remaining two animals were only trained on the medial and lateral positions. For all animals, only reaches when the seed was on the medial and lateral pedestals were included to calculate the success rates. All reaches to all pedestals were included in the neural coding analyses, as they all can be used to study the motor cortical encoding of different kinematic features.

Performance on the task

A successful reach is one where the animal grasped the seed and brought it into the arena. Success rate is defined as the proportion of reaches that are successful out of all reaches made with a seed present on the pedestal in the medial or lateral position (see previous section). Reaches where a seed is not present at the beginning of the reach are not included in the total. Success was scored manually by watching the video recording of each reach (yeah).

Inclusion of reaching data in analysis

Success versus fail: Due to the nature of the task, successful reaches naturally exhibit lower variance across multiple kinematic features compared to failed reaches. For example, the paw must pass within a certain distance of the pedestal during the reach for the animal to grasp the seed. In all neural coding analysis, we include both successful and failed reaches, as both sets of movements can be used to study the motor cortical encoding of different kinematic features.

Timepoints during reaching: We had the least occlusion and the most reliable tracking of the paw and digit markers when the paw was outside of the arena so we focus our analysis of neural coding on the animal's movements there. Hence, only timepoints during reaching were included in the linear model.

One-photon calcium imaging

Image acquisition

One-photon calcium imaging was performed using an Inscopix nVista 3.0 miniscope. During recording sessions, the lightweight (2g) miniscope was attached to a baseplate, which itself was chronically attached to the animal's skull and centered over the chronically implanted grin lens and prism. Data was collected in a single imaging plane at 30Hz. Field of view corresponds to approximately 900 x 650 micrometers, captured in 1236 x 1040 pixels. Each recording session lasted approximately 40 minutes, broken up into 3 to 5 minute recording chunks, separated by breaks of 2 to 4 minutes. Imaging was continuous during a recording chunk, but paused during breaks. When analyzing imaging data from a given session, all data from that session was concatenated and analyzed together.

Locating the same field of view over days

The same field of view was imaged across all days of training. Because the prism is chronically implanted into the neural tissue, the only degree of freedom in identifying the field of view is depth (because of the 90 degree rotation from the prism and its orientation, more "shallow" corresponds to more posterior, whereas deeper into the tissue corresponds to more anterior). On the first day of training, the depth with the most active cells was selected while the animal was moving around the arena but not performing the task. Then, some small blood vessels were identified and used as fiduciary markers to find the same plane on subsequent days. Imaged FOVs are estimated between 50 and 100 micrometers deep in the tissue from the surface of the prism.

Pre-processing of calcium imaging data

Data was pre-processed using Inscopix's proprietary software (Inscopix Data Processing Software, IDPS), including downsampling by averaging (4x downsampling means a 4x4 square of pixels is averaged to one pixel), spatial filtering, and motion correction. We exported the resulting stacks of images for cell identification using Caiman CNMFE (Zhou et al., 2018; Giovannucci et al., 2019) to extract spatial footprints and single-cell fluorescence time series. We took the inferred spike data (S output from Caiman), thresholded at 1 standard deviation above the mean for each neuron on each day separately, then convolved this signal with a gaussian of 30ms standard deviation. This is the neural activity signal used for all analysis.

Registering cells across days

We registered cells across days using CellReg (Sheintuch et al., 2017), with a strict, fixed requirement that across two consecutive days, weighted spatial footprints had to have a spatial correlation of 0.9 to be considered the same cell.

Quantification and statistical analysis

Unless otherwise stated, numbers refer to the following. For description and analysis of behavior, n refers to the number of mice studied, generally n = 5 mice. For description and analysis of neural activity, n refers to the number of neurons recorded from a field of view on one day in one mouse. For comparison of neural encoding properties across days, n refers to the number of pairwise across-day comparisons made, totaled across all cells found on two days that interval of days apart; note that most cells are found on more than two days and thus have more than one comparison made for them.

Analyses were performed on all reaching data, regardless of the success of the reach. See section on kinematic reconstruction for definition of reaching.

Peri-event time histograms (PETHs)

We calculated PETHs from snippets of neural activity where we've locked t = 0 to the start of the reach or the start of the first reach in a bout. We focused on the time window from 100ms before reach onset to 500ms after reach onset. For visualization and for comparison across days, PETHs were normalized to the lowest value = 0, highest value = 1 within this time window.

Linear encoding models of single neurons

On each day separately, we fit a linear model where we predicted the activity of each cell based on measured behavioral and task variables. We used an instantaneous encoding model; neural activity at time t was predicted from behavior and task variables at time t. No time-shifted versions of any predictors were included. Behavioral variables were all interpolated to 30Hz (see kinematic processing section).

Preparing predictor sets/design matrices:

Treatment of predictors: The predictors included are the position and velocity in x and y of the paw and the head (8 variables), the first four PCs of the paw markers, and the first two PCs of the head markers, and time within the session (not including breaks between recording chunks; time variable is equivalent to the index of the calcium imaging data). Only timepoints during reaching (when the paw was outside the arena) were used in the model. Because we want to compare the encoding for the same absolute values of the predictors, we Z-score on all days together rather than on each day separately. For each mouse separately, we treated each predictor as follows: We concatenated all reach timepoints for that predictor across all five days, then z-scored. We then separated the data back into days and performed the covariate balancing procedure described below.

Covariate balancing across days to allow for direct comparison of linear fits: In order to directly compare properties of our linear model fits across days, we must ensure that each predictor has a similar distribution to itself over days, and also that the joint distribution of predictors is similar. Without this step, we cannot guarantee that observed instability is not a result of the animals exploring a slightly different part of kinematic space each day. We use a greedy matching algorithm to balance the joint distribution of all behavioral predictors (the time predictor is not included). **All single day results presented in this paper are covariate balanced across all five days for that animal.** After covariate balancing, each predictor is again z-scored within day before fitting the models.

Train-test split

We used a train-test split of 70% train, 30% test. We did the following procedure to ensure that each train-test split contained a balanced sampling of timepoints from across the session. For all the timepoints included in a day for a given mouse, we split the data into 100 chunks. We then grouped every 10th chunk together, resulting in 10 sets. For example, chunks 1, 11, 21, etc. went into set 1, while chunks 2, 12, 22, etc went into set 2. We then chose 7 random sets to use as the train set, and held out 3 chunks as the test set. We repeated this procedure for 10 cross-validations for all model fits. All linear fit values are reported as averages across the 10 cross-validations. Note that due to the train-test split, r-squared values can sometimes be negative. Within a day, the train-test splits for each cross-validation were the same for all cells.

Fitting single trial models

Single trial models refers to the standard instantaneous linear model fit where we predict neural activity at time t from behavioral and task variables at time t. We fit a family of models for each cell. Full models were fit using all predictors (eq #). We also fit a single variable model for each predictor that included only that predictor (eq #). We also fit a leave one out model for each predictor that included all predictors except for the one in question (eq #). Delta-r-squared values are computed as the drop in r-squared from the full model to the leave one out model for the predictor in question, as schematized in Fig 5a.

All reported r-squared and delta-r-squared values are averaged across 10 cross-validations of the train-test split. Delta-r-squareds are calculated within each cross-validation, then averaged across cross-validations. Due to the train-test split, the Delta R-squared (leave one out models) can sometimes be positive, indicating that removing the predictor in question improved the test R-squared values. This typically occurs for small R-squared values or poor fits, where the model is fitting to noise in the train set and thus cannot generalize to the test set.

Significance testing of full model r-squared values

To test whether a full model fit was better than expected by chance, we did a significance test for each neuron on each day. We shuffled the timepoints of the neural activity vector (predicted variable) and refit the full model. We repeated this procedure 1000 times. If the observed r-squared value was greater than 95% of time-shuffled r-squared values, we deemed the fit for that neuron on that day to be significant.

Fitting trial average models

For each predictor separately, we calculated a time-varying average movement during reaching, locked to the onset of reach. This is very similar to calculating a PETH, but where we average across snippets of a kinematic variable, instead of snippets of neural activity. We built a new design matrix where for each predictor we replaced each reach with a copy of that predictor's average. We only replaced reach timepoints: for a reach that is 10 timepoints long, we replaced those 10 timepoints with the first 10 timepoints of the reach-locked average, whereas for a reach that is 20 timepoints long, we replaced those 20 timepoints with the first 20 timepoints of the reach-locked average. Only reach timepoints were used for the models, as for the single trial models. All trial average models were fit using the same train-test procedure described for single trial models above.

Cosine similarity of r-squared values and coefficients

To quantify the similarity of the tuning profile of a given cell across two days, we compared two sets of linear model fit components: r-squared values and coefficients. For each set separately, we did the following comparison. For each cell, we make a vector out of all of the r-squared values for all models: the cross-validated r-squared and delta-r-squared for all predictor variables and the r-squared for the full model (13 predictors*2 + 1 r-squared for full model = 27 dimensional vector). We then calculate the cosine of the angle between this vector for the cell on day 1 and this vector for the cell on day 2. Because of our test-train split, any r-squared value can be positive or negative, so these vectors are not guaranteed to occupy the same hyper-quadrant. Thus, the cosine of the angle varies between -1 (anti-parallel) and 1 (parallel). We also perform the same procedure for the coefficients from all the single variable predictors for a given cell. The coefficients are unbounded, so the cosine of the angle between possible coefficient vectors also varies between -1 and 1. This approach to quantifying similarity of tuning properties ignores the absolute magnitude of tuning for each predictor and instead compares the magnitude of the r-squared values for the predictors relative to each other.

Control shuffle: Resampled shuffle control built by generating pseudo-vectors where each coefficient is drawn from the pool of coefficients observed across all neurons for that predictor within that day. Coefficients within the same pseudo-vector are drawn independently. The resampled shuffle distribution contains the same number of comparisons for each of the 10 possible pairs of the 5 days of study (5000 total, 500 per day pair). For each specific interval between days, we generated a resampled control distribution. If the observed cosine similarity exceeded 95% of this distribution, the individual comparison was considered significant. It's important to note that using one neuron in the population as a control for another is not an appropriate test because most cells are correlated with each other during motor behavior. Therefore, this analysis focused on the number of cells that showed significant similarities in their encoding profiles across different interval. Note this is comparisons, not cells - an individual cell may show up in multiple comparisons for the same interval, e.g. 1-day interval for a cell that shows up on day 1, 2, and 3.