Pan-cortical 2-photon mesoscopic imaging and neurobehavioral alignment in awake, behaving mice

Evan D. Vickers[1] and David A. McCormick[1,2,3]

[1]Institute of Neuroscience, University of Oregon, Eugene, OR. [2]Professor, Department of Biology. [3]Director, Institute of Neuroscience

Abstract

The flow of neural activity across the neocortex during active sensory discrimination is constrained by task-specific cognitive demands, movements, and internal states. During behavior, the brain appears to sample from a broad repertoire of activation motifs. Understanding how these patterns of local and global activity are selected in relation to both spontaneous and task-dependent behavior requires in-depth study of densely sampled activity at single neuron resolution across large regions of cortex. In a significant advance toward this goal, we developed procedures to record mesoscale 2-photon Ca^{2+} imaging data from two novel *in vivo* preparations that, between them, allow simultaneous access to nearly all of the mouse dorsal and lateral neocortex. As a proof of principle, we aligned neural activity with both behavioral primitives and high-level motifs to reveal the existence of large populations of neurons that coordinated their activity across cortical areas with spontaneous changes in movement and/or arousal. The methods we detail here facilitate the identification and exploration of widespread, spatially heterogeneous neural ensembles whose activity is related to diverse aspects of behavior.

Introduction

Recent advances in large-scale neural recording technology, such as widefield imaging (Musall et al, 2019; Gallero-Salas et al 2021; Esmaeili et al, 2021), large field-of-view (FOV) 2-photon (2p) imaging (Sofroniew et al, 2016; Stringer et al, 2019; Yu et al, 2021), and Neuropixels high-density extracellular electrophysiology recordings (Jun et al, 2017; Steinmetz et al, 2019) have allowed for rapid advancement in our understanding of the relationships between brain-wide neural activity and both spontaneous and task-engaged behavior in mice. For example, these techniques can now be deployed in recently developed behavioral paradigms that allow for the separation of decision, choice, and response periods in 2-alternative forced choice (2-AFC) discrimination tasks (Guo and Svoboda, 2014), the standardization of training and performance analysis across laboratories (The International Brain Laboratory et al, 2021; eLife), and the separation of context-dependent rule representation and choice in a working memory task (Wu and Shadlen, 2020).

One of the striking features of neocortical neuronal activity is how strongly changes in behavioral state, such as task engagement, movement, or arousal, affect the spontaneous and evoked activity of neurons within visual, auditory, somatosensory, motor and other cortical

regions (Niell and Stryker, 2010; McGinley et al, 2015; Musall et al, 2019; Stringer et al, 2019). However, this is not to say that these effects are uniform across the cerebral cortex (Morandell et al, 2023; Wang et al, 2023). Neurons exhibit diversity in the dependence of their activities on arousal and behavioral state both between and within cortical areas (Niell and Stryker, 2010; McGinley et al, 2015; Shimoaka et al, 2018), and these areas are active at different times during a rewarded task (Salkoff et al, 2020; Esmaeili et al, 2021; Gallero-Salas et al 2021). Furthermore, the arousal dependence of membrane potential across cortical areas has been shown to be diverse and predictable by a temporally filtered readout of pupil diameter and walking speed (Shimoaka et al, 2018). This makes simultaneous recording of multiple cortical areas essential for comparison of the dependence of their neural activity on arousal/movement, because combining multiple recording sessions with pupil dilations and walking bouts of different durations will necessarily obscure the true spatial correlation structures.

Areas involved in sensory decision making are often far from each other (Gallero-Salas et al, 2021) and can exhibit coordinated state-dependent changes in functional coupling (Clancy et al, 2019). Also, multimodal sensory information is multiplexed and combined as it ascends across the cortical hierarchy (Coen et al, 2023). For these reasons, understanding the brain activity underlying optimal performance during multimodal, task-engaged behavior will require dense sampling of many brain areas at single neuron resolution across lateral, dorsal, and frontal cortices simultaneously at a temporal resolution high enough to describe both spontaneous behavioral state transitions and the neural dynamics relevant for a given task.

Dense intra-cortical sampling at a fixed depth/cortical layer across many areas is not possible with current Neuropixels probes (Steinmetz et al, 2021), 1-photon widefield imaging can be contaminated by neuropil and hemodynamic signal (Waters, 2020; Valley et al, 2021) and does not achieve single cell resolution (but see Yoshida et al, 2018; Kauvar et al, 2020), and standard 2-photon (2p) imaging is limited by scanning speed (see Gong et al, 2015) and field of view spatial extent (FOV; i.e. simultaneously imageable, either contiguous or non-contiguous, regions; Allen et al, 2017; Hattori et al, 2019; but see Yu et al, 2020). Recent advancements in cranial window preparations have enabled imaging over large portions of the dorsal cortex (Kim et al, 2016; Stringer et al, 2023, bioRxiv). However, simultaneous 2-photon imaging over a large cortical area including both dorsal and lateral regions, particularly in a preparation that allows for simultaneous imaging of the major primary sensory cortices (auditory, visual, and somatosensory) and frontal choice/motor areas (M1, M2) in awake, behaving mice has not been previously shown.

Here, we developed a set of new techniques and integrated them with existing technologies in order to overcome the limitations of current state-of-the-art methods and provide the first pan-cortical 2p assays at single-cell resolution in awake, behaving mice. To achieve this, we designed custom 3D-printed titanium headposts, mounting devices, adapters, and cranial windows, and modified ("Crystal Skull"; Kim et al, 2016; Figs. 1a, d, S1c upper, e) or developed ("A1/V1/M2" or "temporo-parietal"; Figs. 1b, e, S1a, f) two *in vivo* surgical preparations, which we will henceforth refer to as the "dorsal mount" and "side mount", respectively. These two novel preparations enabled mesoscale 2-photon imaging (2p-RAM mesoscope, Thorlabs; Sofroniew et al, 2016) of up to ~7500 individual neurons simultaneously at ~3 Hz across a 5 x 5 mm FOV (Figs. 2, 3, S2; Suppl movies S4, S5; Protocol 1), or up to 800 neurons combined across four 660 x 660 µm FOVs at ~10 Hz (Fig. S2d, e; Protocol 1; see Methods).

In addition, we designed a custom, LabView-controlled behavioral setup with up to 3 high-speed body and face cameras (Figs. 1a, b; Suppl movies S1), allowing us to compare widespread, densely sampled, high-resolution neural activity to movement and behavioral arousal state variation in head-fixed, awake, ambulatory (i.e. running atop a cylindrical wheel) mice under conditions of spontaneous behavior (Figs. 4-6, S3-S7), passive sensory stimulation (Figs. 2, 3, S2), or 2-alternative forced choice (2-AFC) task engagement (Figs. 1c, Protocol 1). We present here detailed methods, along with example recording sessions during spontaneous behavior from both dorsal and side mount preparations, to demonstrate the feasibility of widespread 2-photon cortical neuronal imaging simultaneously with behavioral monitoring.

Methods

Mice. To image cortical activity, we used both male and female, 8-50 week old, CaMKII-tTA x tetO-GCaMP6s (provided by Cris Niell, University of Oregon), CaMKII-Cre x Ai148 (GCaMP6f), and CaMKII-Cre x Ai162 (GCaMP6s) mice (Jackson Labs, Allen Institute). Neural activity was typically imageable between 30 and 180 days post-windowing. Habituation to the experimental setup was performed for 2-3 days prior to 2p data acquisition during spontaneous behavior sessions.

Surgeries. See Protocols II and III, and Supplementary Methods and Materials.

Behavioral videography. See Supplementary Methods and Materials, Overall Workflow, Protocol I. See Supplementary Movies S1.

Widefield imaging and multimodal mapping. See Supplementary Methods and Materials, Overall Workflow, Protocol I, and Multimodal Mapping. See Supplementary Movies S2 and S3.

2-photon imaging and ROI extraction. See Supplementary Methods and Materials, ScanImage 2-photon acquisition. 2p imaging was performed with a Spectra Physics Mai Tai HP-244 tunable femtosecond laser and the Thorlabs Multiphoton Mesoscope. We controlled imaging with ScanImage (2018-2021; Vidrio) running in Matlab, with GPU-enabled online z-axis motion correction. Behavioral synchronization was achieved by triggering 2p acquisition with the first right body camera frameclock signal. Left and posterior cameras were triggered by the same signal, but were subject to variable LabView runtime-related delays on the order of up to ~10 ms and, occasionally, dropped frames. Exposure times for both left and right body cameras, as well as frame-clock times for all three body cameras, were recorded in Spike2 (CED Power1401) as continuous waveforms and events, respectively, to allow for temporal alignment of recorded videos. We used the ScanImage function SI render to combine image strips from large field-of-view (FOV) sessions, and in some cases we used custom Suite2p Matlab scripts (from Carsen Stringer) to align small FOVs from the same or different z-planes in a single session. Rigid and non-rigid motion correction were performed in Suite2p, along with region of interest (ROI) detection and classifier screening of ROIs likely to be neurons. Fluorescence intensity was calculated as changes in fluorescence divided by baseline fluorescence (dF/F), where F was calculated, pixel by pixel, using the ~30-60 s (100-200 frame) rolling 10th or 15th percentile baseline of the neuropil subtracted mean ROI pixel intensity (F-0.7*Fneu). Traces of dF/F were then truncated to match the length of acquired behavioral data and resampled (python: scipy.signal.resample for upsampling, or numpy : slicing for downsampling), along with behavioral data, to 10 Hz for further analysis. ROIs with a Suite2p classifier-assigned cell probability less than 0.5 were excluded from further analysis.

Pose-tracking analysis (DeepLabCut). An example DeepLabCut (DLC) (https://github.com/DeepLabCut/DeepLabCut; Mathis et al, 2018) annotated video with 66 labeled points is shown in Suppl Movie S1 (bottom right). Video snippets were concatenated, cropped, truncated, and aligned to 2p frames with FFMPEG under the Linux subsystem for Windows (Ubuntu 20.04; Windows 10 Enterprise 2004) or Ubuntu 18.04 running on local

network GPU clusters. In the example session shown in Figs. 5, 6 and S7, we analyzed the right Teledyne Dalsa camera video (30 Hz, 1020 x 760 pixel with 2x2 binning, 1” x 1.8” sensor, 90 min video, ffmpeg crop compression to 525.12 MB), which was natively aligned to the 2p acquisition by a common trigger with frameclock and exposure times recorded by a CED 1401 data acquisition device in Spike2 software (Cambridge Electronic Design), running 650K iterations with the resnet_50 model in DLC (see labeled video, Suppl Movie S1, side mount).

Behavioral motif analysis (BSOiD). Pose-tracking data outputs from DLC in the form of comma-separated value files were loaded into B-SOiD (https://github.com/YttriLab/B-SOID; Hsu and Yttri, 2021) along with concatenated, cropped, and spatially downsampled (FFMPEG) right-camera videos (~20-90 minutes, ~100 MB to 1 GB per video). B-SOiD uniform manifold approximation and projection (UMAP) embedding training was performed on between 1 and 4 sessions from the same mouse. Between 20 and 60 labeling space dimensions were typically assigned per session, with between 75 and 98% of features confidently assigned to between ~2-15 motifs per session. Minimum bout time was set at 900 ms, and minimum cumulative motif time per session was set at ~1-5 minutes.

Pixel-wise analysis of cumulative motion energy showed that B-SOiD-identified motifs differed in terms of body-region localized basic movement patterns in a manner that was consistent with human-annotated categorization labels. Thresholding of cumulative motion energy by z-score further confirmed this and made the following kinematic patterns evident: i) In asymmetric pose motifs, a clear region of low movement (dark pixels) is visible in the chest region between the left and right shoulders that is either absent or not as prominent in high-arousal and symmetric pose motifs); ii) Enhanced nose and mouth movement was associated with both high arousal walking and non-walking states; and iii) Increased whisker movement was associated with both high arousal walking and non-walking states.

Results

Recent work with large-scale imaging (e.g. widefield and mesoscale 2p) and electrophysiology (e.g. Neuropixels) recording technologies has suggested that a significant percentage of the variance of neural activity across neocortex is accounted for by rapid spontaneous fluctuations in arousal and self-directed movement during both spontaneous behavior and task performance (Musall et al, 2019; Steinmetz et al, 2019; Stringer et al, 2019; Jacobs et al, 2020; Salkoff et al, 2020; Stringer et al, 2023, bioRxiv). Given that arousal state- and movement-related activity

appears to be a ubiquitous feature of cortical activity across all regions, it would be advantageous to develop methodologies that allow for simultaneous, single neuron resolution, contiguous monitoring of neuronal activity during second-to-second movements and changes in arousal, within both spontaneous and trained behaviors. Here, we harness new imaging and analysis technologies in order to both address this methodological gap and provide a proof of principle test of these methods by examining the relationship of behavioral arousal and movement to detailed spatial patterns of neural activity across the dorsal and lateral neocortex.

Large field of view 2-photon imaging in behaving mice. In order to simultaneously monitor the activity of cortical neurons across a large area (~25 mm^2, or up to 36 mm^2 in some cases) of the dorsal cortex in awake, behaving mice, we chose 2-photon imaging of GCaMP6s fluorescence using existing 2-photon random access mesoscope (Thorlabs) imaging technology (2p-RAM; Sofroniew et al, 2016). We chose this form of data collection for use in head-fixed mice because it allows for: 1) rapid scanning (i.e. it uses a resonant scanner coupled to a virtually conjugated galvo pair) over a large (5x5 mm fully corrected, or up to 6x6 mm imageable), field-curvature corrected field of view (FOV); 2) subcellular resolution in the z-axis to avoid region-of-interest (ROI) contamination by neuropil and neighboring cells (0.61 x 0.61 x 4.25 µm xyz point-spread function at 970 nm laser excitation wavelength); 3) correction for aberrations at all wavelengths between 900-1070 nm, to allow for combined imaging of multiple fluorophores. We achieved all three of these objectives in head-fixed mice that are ambulatory, able to clearly view visual monitors on both sides, and readily able to lick two water spouts (left, right) for 2-alternative forced choice responses, with left, right, and rear videography to monitor each mouse's face, pupil and body.

The neuronal imaging methodology we chose (see below) is not the only one available. For example, another promising and recently developed alternative implementation, Diesel 2p, allows for similarly flexible and large FOV imaging at single cell resolution, but with a significantly higher z-axis point-spread function (~8-10 µm; Yu and Smith, 2021). Despite this drawback, its use of dual scan engines allows for true simultaneous imaging of different cortical areas, while the 2p RAM mesoscope (Thorlabs) only accomplishes this by rapid jumping (~1 ms) between FOVs. Both systems offer excellent corrected field curvatures on the order of +/- 25 µm over ~5 mm.

Several primary design constraints needed to be met to achieve our goal of monitoring neuronal activity over a large portion of the mouse dorsal cortex. These ranged from aspects of the surgical preparation, mounting, and imaging, to neural recording (Figs. 1, S1, Suppl Movies

S1, Protocols II-III). To increase the extent of the neocortex over which we could monitor neural activity, we developed two distinct headpost/cranial window preparations. In the first, the dorsal mount, the mouse is head-fixed upright on a cylindrical running wheel and the objective is vertical (Fig. 1a). This allows for monitoring of neural activity over large regions of both cerebral hemispheres, from the posterior aspects of visual cortex to the anterior aspects of motor cortex, and laterally to the dorsal-most aspects of auditory cortex, depending on the placement of the 2p mesoscope objective (Fig. 1d). The second preparation, the side mount, required the head of the mouse to be rotated 22.5^{o} to the left (or right, not shown), with the microscope objective vertical or tilted 1-5 degrees to the right, so as to extend the 5x5 mm imaging FOV to include the auditory cortex and other neighboring ventral/lateral cortical areas (Fig. 1b, e).

We observed that mice adapt well to head tilt in the side mount preparation, and will readily whisk, walk or run, and learn to perform lick response tasks in this configuration. Keeping the microscope objective in a vertical or near-vertical orientation, and rotating the mouse's head (instead of rotating the objective), significantly improved the manageability of the water meniscus between the objective and the cranial window. However, if needed, the objective of the Thorlabs mesoscope may be rotated laterally up to +/- 20° for direct access to more ventral cortical areas, for example if one wants to use a smaller, flat cortical window that requires the objective to be positioned orthogonally to the target region.

Imaging in this novel side mount preparation was restricted to one hemisphere and a thin medial strip (~1 mm wide) of the contralateral hemisphere. This allowed for imaging from roughly the anterio-medial areas of V1 to medial and medio-lateral secondary regions of motor cortex (or slightly farther, if the anterior-posterior axis of the brain is oriented along the diagonal of the 5 x 5 mm FOV). This anterior-posterior reach was similar to that observed in the dorsal mount preparation. However, the side mount windowing and the accompanying mounting procedure significantly increased our ability to image neural activity in lateral (ventral) cortical aspects, including auditory, somatosensory, and association cortical regions (Fig. 1e), notably without the need for substantial rotation of the objective.

To achieve 2p neuronal imaging from broad cortical regions, we created large and stable, custom (designed in AutoDesk Inventor) 3D-printed titanium (laser sintered powder with active cooling and shot-peened post-processing) headposts (Figs. 1d, 1e, S1e, S1f, Suppl Design Files; i.materialise.com and sculpteo.com) and glass (0.21 mm thick, Schott D263T) cranial windows (Figs. 1d, 1e, S1a, c, d, e, and f; Suppl Movies S1; Suppl Design files: https://github.com/vickerse1/mesoscope_spontaneous; Labmaker.org for dorsal mount; TLC International for cutting, and GlasWerk for bending, for the side mount and early prototypes of

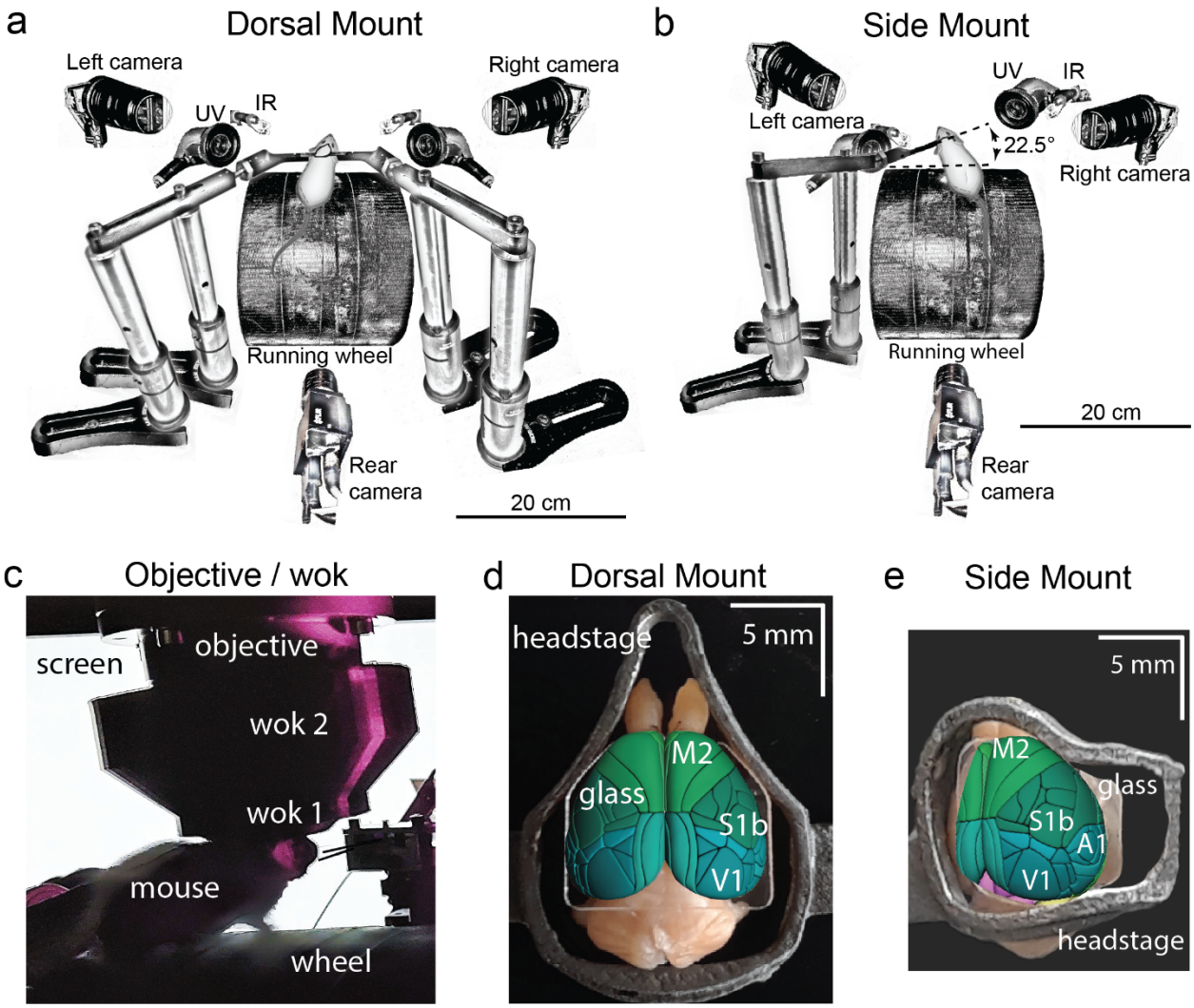

Figure 1. Technical adaptations for dual-mount *in vivo* pan-cortical imaging with the Thorlabs 2p-RAM mesoscope. a) Mesoscope behavioral apparatus for the dorsal mount preparation. Mouse is mounted upright on a running wheel with headpost fixed to dual adjustable support arms mounted on vertical posts (1" diameter). Behavior cameras are fixed at front left, front right, and center posterior, with ultraviolet and infrared light-emitting diodes aligned with goose-neck supports parallel to right and left cameras. **b)** Mesoscope behavioral apparatus for side mount preparation. Same as in (a), except that the mouse is rotated 22.5 degrees to its left so that the objective lens (at angle 0) is orthogonal to its center-of-mass of preparation in the right cortical hemisphere. The objective can rotate +/- 20 degrees medio-laterally, if needed, to optimize imaging of any portion of the cortex under the cranial window. The right behavior camera is positioned more posterior and lower than in a), to allow for imaging of the eye under the acute angle formed with the horizontal light shield, shown in c). **c)** Mouse running on wheel with side mount preparation receiving visual stimulation from an LED screen positioned on the left side, with linear motor-positioned dual lick-spouts in place and 3D printed vertical light shield (wok 2) attached to rim of flat shield (wok 1) to block extraneous light from entering the objective. **d)** Overhead view of dorsal mount preparation with 3D printed titanium headpost, custom cranial window, and Allen CCF aligned to paraformaldehyde-fixed mouse brain. Motor region = light green, somatosensory = dark green, visual = dark blue, retrosplenial = light blue. Olfactory bulbs (anterior) at top, cerebellum (posterior) at bottom of image. Note ridge along perimeter of headpost for fitted horizontal light shield (wok 1) attachment. **e)** Rotated dorsal view (22.5 degrees right) of side mount preparation with 3D printed titanium headpost, custom cranial window, and Allen CCF aligned to paraformaldehyde-fixed mouse brain. Auditory region shown in light blue, ventral and anterior to visual areas, and ventral and posterior to somatosensory areas (right side of image is lateral/ventral, and left side of headpost perimeter is medial/dorsal). Other regions shown with the same color scheme as in d). UV = ultraviolet, IR = infrared, M2 = secondary motor cortex, S1b = primary somatosensory barrel cortex, V1 = primary visual cortex, A1 = primary auditory cortex.

the dorsal mount). We incorporated these custom made headposts and cranial windows into the dorsal mount (Figs. 1a, d, Fig. S1c upper, S1d upper left, S1e), and side mount (Figs. 1b, e, S1a, c lower, d lower left and right, and upper right, S1f) preparations.

To mimic the curvature of the brain and reduce tissue compression, which is a problem with flat coverslips with a diameter larger than ~3 mm, cranial windows were curved by heating pre-cut (TLC International) glass pieces over 9 or 10 mm bend-radius molds (GlasWerk). The bend radius (i.e. radius of half cylinder mold over which the melted glass is bent to achieve its curved shape) was fixed at 10 mm for Labmaker.org dorsal mount windows (Fig. S1e, right; Suppl Movie S1 dorsal mount; early attempts with 11 or 12 mm bend radii failed), or, for windows custom designed in collaboration with GlasWerk (Fig.S1f, right; Suppl Movie S1 side mount), the bend radius was set at either 9 mm, for a tight fit to ventral auditory areas, or at 10 mm, to enable simultaneous imaging of the entire preparation at a single focal depth.

Within a field of view of 5x5 mm, curvature of the brain, especially in the medial-lateral direction, is a significant problem. To partially compensate for brain curvature, online field-curvature correction (FCC) was applied in ScanImage (Vidrio Technologies, MBF Biosciences) during mesoscope image acquisition. In addition, complementary fast-z “sawtooth” or “step” corrections were applied when all ROIs were acquired at the same or different z-planes, respectively (see Sofroniew et al, 2016). Despite this, small uncorrected discrepancies in the depth of the imaging plane below the pial surface may have remained, especially along the long-axis of acquisition ROIs (i.e. mediolateral). It is likely that optical effects of the curvature of the glass partially compensated for these differences in targeted imaging depth, although we did not quantify this effect. In cases where we imaged multiple small ROIs, nominal imaging depth was adjusted in an attempt to maintain a constant relative cortical layer depth (i.e. depth below the pial surface; ~200 m offset due to brain curvature over 2.5 mm mediolateral distance, symmetric across the center axis of the window).

The mouse was restrained by fixation of the headpost with two countersunk screws to either a fixed (Fig. S1b, top) or adjustable aluminum support arm (custom; Figs. 1a, b, S1b, bottom) and mounting apparatus (Figs. 1a, b). For the dorsal mount preparation, two headpost support arms, one on each side of the head, were used (Fig. 1a, S1c, top; outer screw of both headpost wings used) while for the side mount preparation, a single, left side support arm was sufficient (Fig. 1b, S1c, bottom; inner and outer screws in the left headpost wing were both used). It was important to fix the headpost mounting arms to the top of the airtable with 1 inch diameter vertical mounting posts (Thorlabs RS6P8, RSH4, PF175), making sure all joints were clean of debris and tightened securely, to reduce movement artifacts.

Both preparations were found to be relatively free of vibration and movement artifacts and allowed for micro-adjustments (i.e. pitch, yaw, and roll) of the mouse relative to both the objective and the running wheel (Fig. 1a, b, c). Increased positioning flexibility of the support-arm assembly was achieved through either of 2 ways: 1) by positioning it on a distal base-attached ball-joint mount (Thorlabs, SL20 articulating base; not shown); 2) through the use of a custom adjustable headpost support arm that was proximally adjusted around a ball-and-joint assembly (Fig. S1b, bottom). This adjustable support arm operated through the use of a custom wrench and reverse-threaded (left-handed) nut for initial coarse tightening, and dual micro-clamps for rotational stabilization during tightening. Four micro-hex wrench driven set-screws, two on either side of the ball-and-joint fitting, were used to achieve the final, fully locked state for imaging. The adjustable support arm was found to be superior to the base-attached ball joint mount in allowing for micro-positioning of the animal in relation to the running wheel, objective, and lick spouts after initial fixation (mounting) of the headstage to the support bar.

In order to perform 2-photon neuronal imaging, the large 2p-RAM water-immersion objective (~10x net magnification, 0.6 NA, ~1.3 kg, 12 mm diameter tip, 25.6 mm shaft) must be positioned between 2.2 and 2.8 mm (i.e. ~working distance minus window thickness) from the curved glass surface of the cranial window while maintaining a stable water meniscus. Providing a base for the water meniscus while also blocking incident light entry was achieved by creating a custom 3D printed plastic light shield (AutoDesk Inventor, MakerGear M3 printer, PLA) that was attached to the protruding, fitted rim of the headpost with 170 FAST CURE Sylgard ("wok one"; Fig. 1c, S1d). A second light shield ("wok two") served to further block light entry into the objective and was attached to the base of the lower light shield ("wok one") by fitting of a U-shaped profile over a vertically protruding single edge profile along the perimeter of the first, lower light shield ("wok one"; Fig. 1c). Together, these two custom-printed light shields prevented incident light (from the video stimulus monitor, the ultraviolet LEDs used for controlling the baseline and dynamic range of pupil diameter, and the infrared LEDs used to illuminate the mouse) from entering the imaging objective and thereby either contaminating or tripping the photomultiplier tubes (PMTs). Line-of-sight for the animal to the video stimulus monitor was retained, and this design allowed for free vertical and rotational (over a limited range) movement of the objective lens, which was contacted directly only by the water meniscus.

Direct left, right, and posterior camera angles, or "lines of sight", were preserved to enable reliable recording and proper illumination of pupil diameter, whisker pad movement, and the movement of other body parts such as the nose, mouth, ear, paws, and tail (Fig. 1a, b; see

Suppl Movies S1). Adjustable positioning of two conductance-based lick spouts for 2-alternative forced choice (2-AFC) task performance (lick left, lick right) and reward delivery was achieved with a rapidly translatable motorized linear stage (Fig. 1c; 63.7 ms total travel time over 7 mm; Zaber Technologies). This system allowed for rapid withdrawal of lick spouts between trials, and presentation of the lick spouts during the response period of each trial.

Widefield imaging and optogenetic stimulation in the 2p-RAM mesoscope. Examining neuronal activity at the single cell level and relating it to stimulus-evoked activity observed at the widefield level (Figs. 2, 3), as well as manipulating regions of the cortex through localized light delivery and optogenetics (see Suppl Movie S7), required precise optical alignment across all three of these methods. First, a standardized cortical map needed to be fitted onto images of skull landmarks and the cortical vasculature by multimodal sensory widefield mapping (Figs. 2, 3; a,b, c, top). Then, a method was needed for transferring this cortical map onto the field of view of the mesoscope so that each individual neuron could be assigned to a cortical area for subsequent analyses (Figs. 2, 3; c, bottom, d). Finally, the spatial coordinate system of the 1p opto-stimulation laser routed in through an auxiliary light-path of the mesoscope needed to be aligned to that of the 2p laser so that specific cortical subregions could be targeted for optogenetic inhibition by light activated, ChR2-mediated excitation of parvalbumin interneurons in Thy1-RGECO x PV-Cre x Ai32 mice (Suppl Movie S7).

In order to align 2p maps of single neuron activity with functional 1p cortical maps determined through pseudo-widefield imaging (i.e. combined reflected and fluorescence light imaging with a standard format CCD camera), we designed a method for reliably aligning a standardized cortical map to an image of the cortical vasculature pattern, which can be directly visualized with both widefield and 2p imaging techniques for each brain (see Protocol V). The epifluorescence light source (Excelitas) for pseudo-widefield imaging and visualization of vasculature for targeting of the 2p FOV was controllable by a dial, shutter, and remote foot-pedal for optimal ease of positioning the objective in the water meniscus at a distance of ~2.2 mm from the headpost and cranial window (Figs. 1c, and S1d, top and bottom right).

This technique required several modifications of the auxiliary light-paths of the Thorlabs mesoscope. For switchable blue/green widefield imaging and 2p imaging, we used a 469/35 Semrock excitation filter, 466/40 Semrock dichroic (1st, top cube), and mirror (2nd, bottom cube). For this standard configuration, which we used with GCaMP6s mice, an electronically positionable dichroic (T580-160, Chroma) was positioned in the light-path just upstream of the mesoscope objective to enable the user to switch on-the-fly between widefield imaging (i.e.

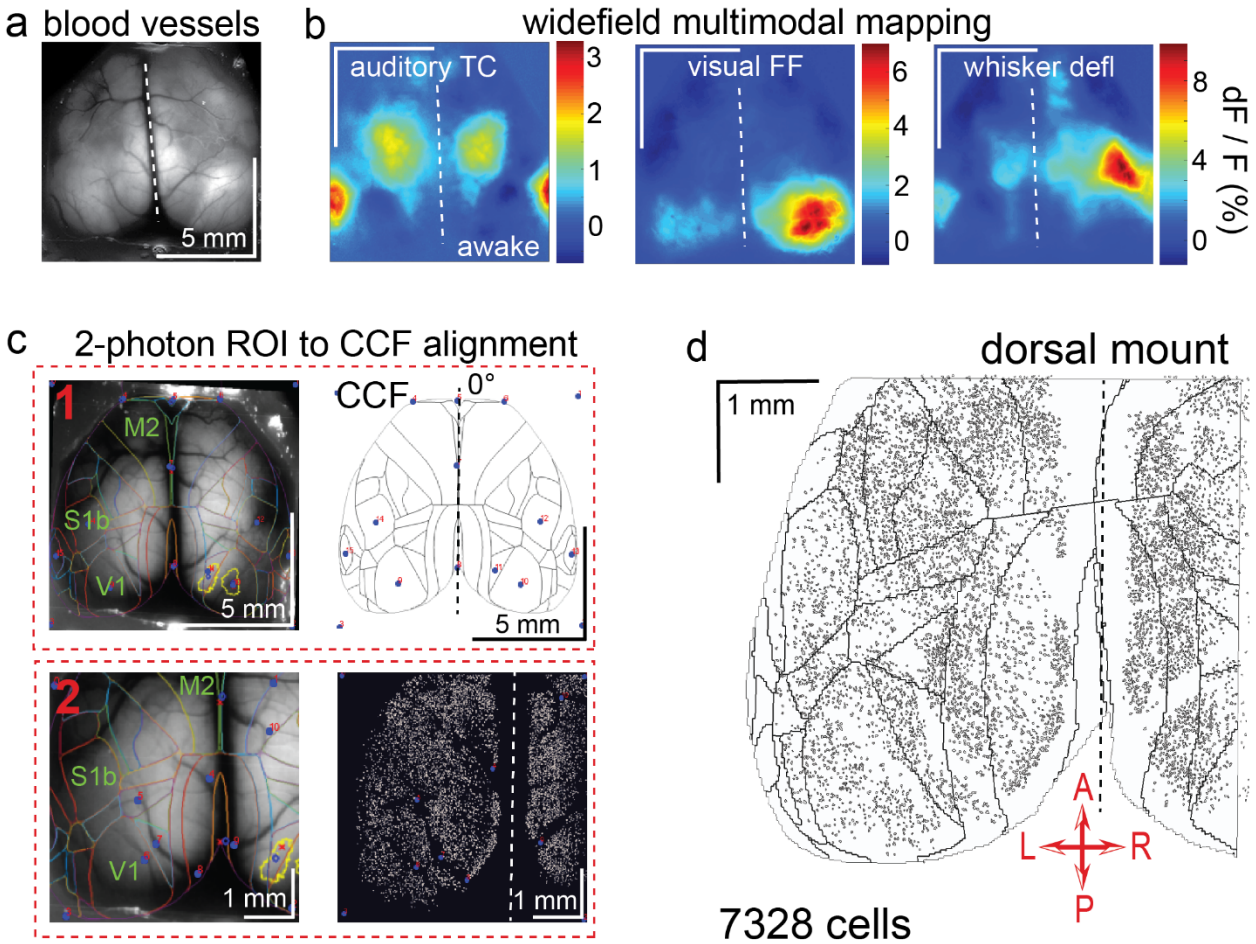

Figure 2. Widefield-imaging based multimodal mapping (MMM) and CCF alignment, ROI detection, and areal assignment of GCaMP6s$^+$ excitatory neurons for the dorsal mount preparation. a) Mean projection of 30 s, 470 nm excitation epifluorescence widefield movie of vasculature performed through a dorsal mount cranial window. Midline indicated by a dashed white line. **b)** Mean 1 s baseline subtracted dF/F images of 1 s responses to ~20-30 repetitions of left-side presentations of a 5-40 kHz auditory tone cloud (auditory TC; left), visual full field (FF) isoluminant Gaussian noise stimulation (center), and 100 ms x 5 Hz burst whisker deflection (right) in an awake dorsal mount preparation mouse. **c)** Example 2-step CCF alignment to dorsal mount preparation, performed in Python on MMM masked blood-vessel image (upper left), rotated outline CCF (upper right), and Suite2p region of interest (ROI) output image containing exact position of all 2p imaged neurons (lower right). Yellow outlines show the area of masks created from the thresholded mean dF/F image for, in this example, repeated full-field visual stimulation. In step 1 (top row), the CCF is transformed and aligned to the MMM image using a series of user-selected points (blue points with red numbered labels, set to matching locations on both left and right images) defining a bounding-box and known anatomical and functional locations. In step 2 (bottom row), the same process is applied to transformation and alignment of Suite2p ROIs onto the MMM using user-selected points defining a bounding box and corresponding to unique, identifiable blood-vessel positions and/or intersections. Finally, a unique CCF area name and number are assigned to each Suite2p ROI (i.e. neuron) by applying the double reverse-transformation from Suite2p cell center location coordinates, to MMM, to CCF. **d)** CCF-aligned Suite2p ROIs from an example dorsal mount preparation with 7328 neurons identified in a single 30 min session from a spontaneously behaving mouse. TC = tone cloud, FF = full-field, defl = deflection, dF/F = change in fluorescence over baseline fluorescence, CCF = Allen common coordinate framework version 3.0, M2 = secondary motor cortex, S1b = primary somatosensory barrel cortex, V1 = primary visual cortex, A = anterior, P = posterior, R = right, L = left.

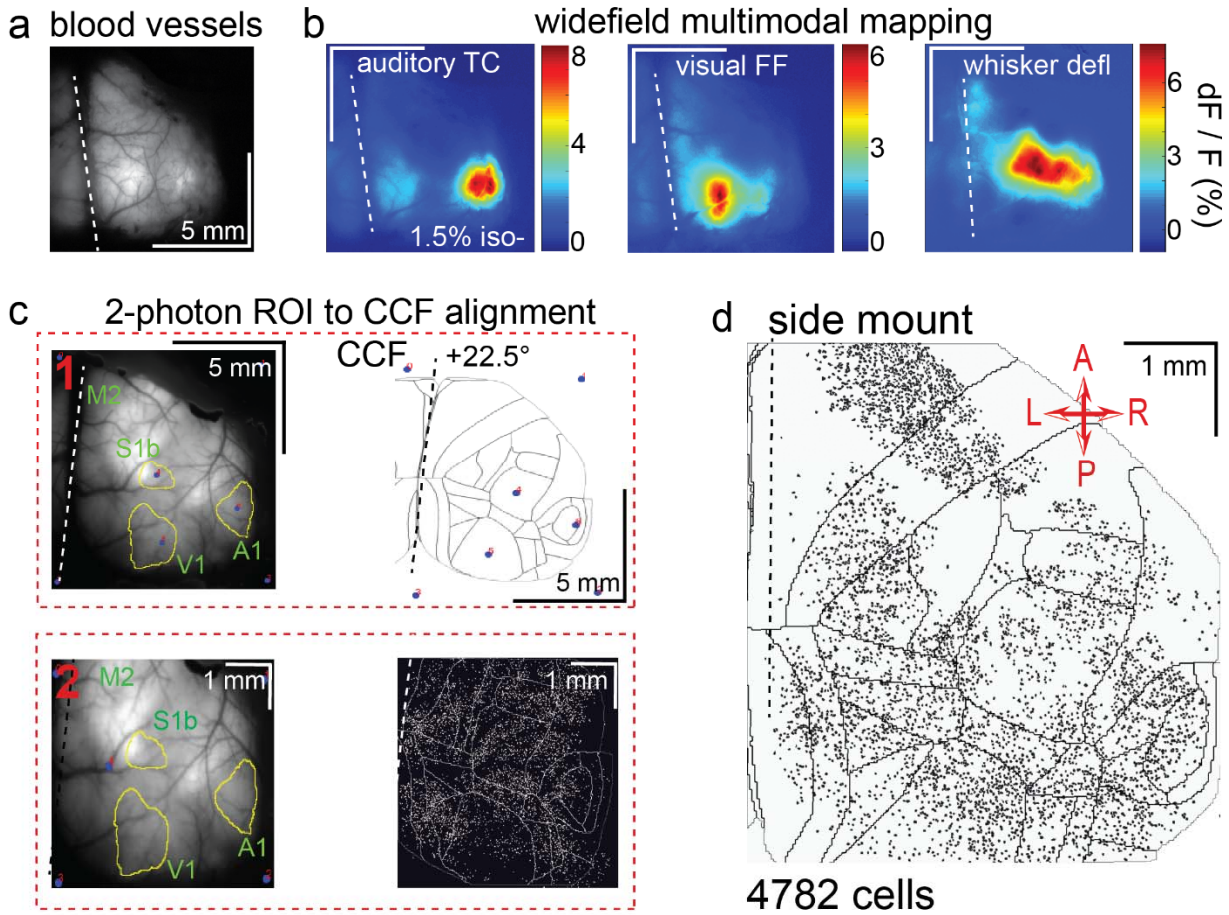

Figure 3. Widefield-imaging based multimodal mapping (MMM) and CCF alignment, ROI detection, and areal assignment of GCaMP6s$^+$ excitatory neurons for the side mount preparation. a) Mean projection of 30 s, 470 nm excitation fluorescence widefield movie of vasculature performed through a side mount cranial window. Midline indicated by a dashed white line. **b)** Mean 1 s baseline subtracted dF/F images of 1 s responses to ~20-30 repetitions of left-side presentations of a 5-40 kHz auditory tone cloud (auditory TC; left), visual full field (FF) isoluminant Gaussian noise stimulation (center), and 100 ms x 5 Hz burst whisker deflection (right) in an example side mount preparation mouse under 1.5% isoflurane anesthesia. **c)** Example 2-step CCF alignment to side mount preparation, performed in Python on MMM masked blood-vessel image (upper left), rotated outline CCF (upper right), and Suite2p region of interest (ROI) output image containing exact position of all 2p imaged neurons (bottom right). Yellow outlines show the area of masks from thresholded mean dF/F image for repeated auditory, full-field visual, and/or whisker stimulation. In step 1 (top row), the CCF is transformed and aligned to the MMM image using a series of user-selected points (blue points with red numbered labels, set to matching locations on both left and right images) defining a bounding-box and known anatomical and functional locations. In step 2 (bottom row), the same process is applied to transformation and alignment of Suite2p ROIs onto the MMM using user-selected points defining a bounding box and corresponding to unique, identifiable blood-vessel positions and/or intersections. Finally, a unique CCF area name and number are assigned to each Suite2p ROI (i.e. neuron) by applying the double reverse-transformation from Suite2p cell center location coordinates, to MMM, to CCF. **d)** CCF-aligned Suite2p ROIs from an example side mount preparation with 4782 neurons identified in a single 70 min session from a mouse performing our 2-alternative forced choice (2-AFC) auditory discrimination task. TC = tone cloud, FF = full-field, defl = deflection, dF/F = change in fluorescence over baseline fluorescence, CCF = Allen common coordinate framework version 3.0, M2 = secondary motor cortex, S1b = primary somatosensory barrel cortex, A1 = primary auditory cortex, V1 = primary visual cortex, A = anterior, P = posterior, R = right, L = left.

dichroic in light path) and stand-alone 2p imaging (i.e. dichroic out of light path). Pseudo-widefield images (ThorCam) and 2p images (ScanImage) were aligned to be parfocal so that the user could switch directly from the vasculature image at a desired CCF location to the 2p image.

For the combination of pseudo-widefield imaging and rapid, targetable optogenetic 1-photon (1p) stimulation with concurrent 2p imaging, which we used with PV-Cre x Ai32 x Thy1-RGECO mice (Suppl Movie S7), we established dual coupling of a broadband fluorescence light-source (Excelitas) and a 473 nm laser driver (SF4C 473, Thorlabs) to converging, dual input auxiliary light paths of the Thorlabs mesoscope via two in-series switchable filter cubes (DFM1T1 cube, Thorlabs). The 473 nm laser driver was connected to a single open loop, high speed buffer (50 LD, Thorlabs), and targeted to the coordinate system of the 2p laser with a grid-calibrated, auxiliary galvo-galvo scanner (GVSM002, Thorlabs). This allowed both orange epifluorescence excitation light and collimated light from the blue 1p laser to reach the mouse via auxiliary light paths, and for 2p imaging and blue 1p laser optogenetic stimulation to occur pseudo-simultaneously, in that an electronically controlled shutter was positioned to block the light path to the PMT and interrupt image acquisition only during the brief presentation of the optogenetic stimulus (~100 ms; see Suppl Movie S7).

Reduction of resonant scanner noise. Resonant scanners in 2p microscopes emit intense sound at the resonant mirror frequency, which is well within the hearing range of mice (~12.5 kHz emitted in the Thorlabs mesoscope). Because scanning precision requires the scanner to remain at a stable, elevated temperature, the scanner must remain on during the entire experimental session (i.e. up to ~2 hrs per mouse). The unattenuated or native high-frequency background noise generated by the resonant scanner causes stress to both mice and experimenters, and can prevent mice from achieving maximum performance in auditory mapping, spontaneous activity sessions, auditory stimulus detection, and auditory discrimination sessions/tasks.

To reduce this acoustic noise, we encased the resonant scanner and attached light path tubes with a custom 3-dimensional (3D)-printed assembly containing dense interior insulating foam (see Suppl Materials: https://github.com/vickerse1/mesoscope_spontaneous; “3D scanning and honeycomb patterned nylon print” (University of Oregon Innovation Disclosure #DIS-23-001, US provisional patent application UOR-145-PROV). It was critical to use 3D scanning of encased components in the design of the noise reduction shield so that the assembly would closely follow all of the surface contours and fit accurately in spaces with tight

tolerances. The result of this encasement was a large reduction in resonance scanner sound, from ~60 dB to ~5 dB measured at the head of the mouse (Bruel and Kjaer ¼" pressure filled microphone 4938-A-011; i.e. below the mouse hearing threshold at 12.5 kHz of roughly 15 dB). By comparison, encasements designed using standard 3D-printing techniques (i.e. not based on a 3D scan) were only able to achieve a noise reduction of ~30 dB, to a level still audible to both mice and humans.

Cortical alignment to the common coordinate framework v3.0 (CCF) map. Interpretation of the diversity of activity of thousands of neurons simultaneously identifiable with the mesoscope requires proper cortical areal localization, including that of areas far from primary sensory cortices. Such mapping is not routinely possible directly on the 2-photon mesoscope when performing neuronal-level imaging, due to both the heterogeneity of responses within primary sensory cortices at the single neuron level, and to animal-to-animal variations in the spatial extent of cortical regions (de Vries et al, 2020; Bimbard et al, 2023). To facilitate the assignment of neurons to cortical areas, we sought to align the Allen Institute Common Coordinate Framework (CCF; Wang et al, 2020) to our 2p neuronal imaging results, using blood vessels, skull landmarks, and widefield imaging responses as intermediaries.

Alignment of the CCF to the dorsal surface of cortex and subsequent image-stack registration based on widefield imaging Ca^{2+} fluorescence data has been performed elsewhere with a variety of techniques. These include the use of skull landmarks (e.g. bregma and lambda; Musall et al, 2019), responses to unimodal sensory stimuli (Gallero-Helmchen, 2021), visual field mapping (Zhuang et al, 2017), and/or autocorrelation maps of spontaneous activity (Peters et al, 2021). Few studies, however, have performed such alignments with rotated cortex (i.e. imaging lateral portions of the cortex at an angle; but, see Esmaeili et al, 2021).

To summarize our approach, we used a technique where, for both the dorsal mount (Figs. 2, S2a) and side mount (Figs. 3, S2b) preparations, we first imaged 1p widefield GCaMP6s responses to unimodal passive sensory stimulation (e.g. visual, auditory, and somatosensory) to create a multimodal map (MMM) consisting of multiple sensory area masks on top of the cortical vasculature. We then overlaid skull landmarks (e.g. bregma and lambda), imaged with reflected green light, onto the MMM. The Allen CCF was then warped onto the vasculature/skull/MMM overlay, and the mean 2p image was warped onto the MMM by vasculature alignment. Finally, each neural ROI in the 2p image was assigned to a position in the MMM x-y coordinate system and a corresponding CCF area based on the final overall

alignment of the MMM, CCF, and 2p image. The following sections describe these steps in more detail.

Widefield multimodal mapping. To create the multimodal map, each mouse was put under light isoflurane anesthesia (~1.5%) after head post surgery, but before implantation of the cranial window, and exposed to 5 minutes each of full-field visual (vertical and horizontal stationary grating patches; 0.16 cpd, 30 deg; Michaiel et al, 2019), tone-cloud auditory (a series of overlapping 30 ms duration tones randomly selected from a frequency range of 5-40 kHz and presented at 100 Hz; Xiong et al, 2015), and piezo-driven whisker deflection (or, in some cases, forelimb and trunk stimulation; see Gallero-Helmchen, 2021). For whisker deflection, a 5 Hz burst of 100 ms duration was used, consisting of posterior to anterior sweeps. The whisker deflector was a custom 3D-printed triangular polylactic acid (PLA) piece mounted on a 21-gauge needle and attached to a PL140.11 piezo-actuator with epoxy glue. It was actuated with a Physik-Instrumente controller driven by a custom pulse sequence generated in Spike2 and delivered from an analog output of a CED Power 1401, with an inter-stimulus interval of ~10 s (Figs. 2b, right, and 3b, right; see also Suppl Movies S2 and S3).

Averaged, baseline-subtracted dF/F responses (Figs. 2b, 3b) were then thresholded, masked, outlined, and layered onto the blood vessel image (Figs. 2a, c, 3a, c, S2a, b) along with bregma and lambda skull landmarks (Fig. S1a) from 530 nm (green) reflected-light skull-imaging using a custom protocol/macro in Fiji (ImageJ; Figs. 2c, 3c, S2a, b). This multimodal mapping procedure and alignment to the CCF and skull landmarks was repeated if significant changes in vasculature pattern occurred over the days/weeks of the experiment. More precise maps of visual cortical areas can be achieved through visual field mapping using a topographic stimulus consisting of a bar sweeping in azimuth or elevation (Garrett et al, 2014). This technique was not applied here, because our goal was global alignment of the Allen Institute cortex-wide CCF to our dorsal or side mount views, and not precise alignment to visual subareas *per se*.

Co-alignment of 2p image to vasculature and CCF. The MMM was aligned to an overlay of the rotated CCF edges and region of interest (ROI) masks using custom Python code and the built-in function “PiecewiseAffineTransform”, given a user-supplied series of bounding-box and alignment points common to both images (Figs. 2c (upper, part 1), 3c (upper, part 1)). A second, similar alignment was then performed between the multimodal map and all neural ROI locations (i.e. Suite2p-identified neurons) relative to the 2p image plane; here, vasculature is inferred from

the pattern of gaps (e.g. vascular “shadows”) in the spatial distribution of neurons (Figs. 2c (lower, part 2), 3c (lower, part 2)), and can be confirmed by parfocal pseudo-widefield (i.e. combined reflected and epifluorescence light) imaging directly on the 2p Thorlabs mesoscope. Note that in 2p imaging below the pial surface, the effective/apparent width of the vasculature, or its “shadow”, is significantly larger than that of the actual blood vessel and its width increases with depth. The transforms resulting from each of these two alignments were then applied in serial (i.e. multimodal map to CCF, then multimodal map to 2p neural image; Figs. 2c and 3c) to overlay the neural image directly onto the cortical map, using outlines, and to assign a unique CCF area identifier (i.e. name and ID number) to each Suite2p-extracted, 2p-imaged neuron using CCF masks (Figs. 2d, 3d).

2-photon imaging across broad regions of the dorsal and lateral cortex in behaving mice. Previous 2-photon imaging studies have demonstrated that arousal and/or orofacial/body movement can explain a significant proportion of the variance in spontaneous and sensory evoked neural responses in visual and other restricted dorsal cortical regions (Musall et al, 2019; Stringer et al, 2019). Here, as a proof of principle, we examined the generality of this finding by simultaneously monitoring neuronal activity across broad regions of the dorsal and lateral cortex with our dorsal and side mount preparations. Previous studies suggest that there may be both commonalities as well as heterogeneity in the effects of changes in arousal/movement on spontaneous and sensory-evoked responses in different cortical areas (McGinley et al, 2015; Musall et al, 2019; Stringer et al, 2019). For example, running typically enhances the gain of visually evoked responses in the mouse primary visual cortex (Niell and Stryker, 2010), while it significantly decreases that of evoked auditory responses in primary auditory cortex (Zhou et al, 2014; McGinley et al, 2015).

Assessing such potential differences between functionally distinct, and potentially distant, cortical areas requires simultaneous imaging across multiple cortical regions while maintaining adequate imaging speed, quality, and resolution. To achieve this level of 2p imaging across several millimeters of cerebral cortex (e.g. from visual to motor or auditory cortical areas), we first needed to optimize the parameters of our mesoscope imaging methods and protocols.

Conventional 2-photon imaging using a preparation similar to our dorsal mount (“crystal skull”; Kim et al., 2016) previously employed serial acquisition of ~1 x 1 mm FOVs (Kim et al, 2016). Other 2p Thorlabs mesoscope imaging studies have used acquisition protocols targeting z-stacks of 600 x 600 µm FOVs in barrel cortex (3 planes at 7 Hz, 1.17 µm/pixel; Peron et al,

2015), 900 x 935 µm FOVs in visual cortex (11 planes at ~1 Hz; Stringer et al, 2019), and three adjacent 600 x 1800 µm FOVs in CA1 hippocampus (Sun et al, 2023; bioRxiv). Here, with our dorsal and side mount preparations, we were able to routinely image between 2000 and 7600 neurons per session over up to ~25 mm^2 of dorsal cortex simultaneously at ~3 Hz with a resolution of 5 µm/pixel in GCaMP6s mice (combined total in both preparations including all large-FOV sessions: N=17 mice, n=91 sessions, ~350,000 neurons; Fig. S2e, Suppl Movie S4).

In order to achieve spatially broad and dense sampling, while monitoring changing activity in each neuron as frequently as possible, we typically tiled 7 mediolaterally-aligned 5000 x 660 µm FOVs (i.e. mechanical scanning along a long-axis oriented mediolaterally, combined with fast resonant-scanning across multiple short-axes oriented anterior-posterior) over either posterior-medial, posterior-lateral, or anterior dorsal, cortex at a pial depth of ~200-300 µm (cortical layers 2/3). Bidirectional scanning (with a scan-phase correction of between 0.9-1.0) and field-curvature correction were typically enabled, with a 1 ms “flyback” and “frame” time, which are defined in ScanImage as the time delay for retargeting of the laser from the end of one line or frame acquisition to the beginning of the next, respectively. Here, there is a trade-off between longer flyback times, which reduce imaging speed but minimize positioning errors at the beginning of each scan line or ROI, and shorter values, which accelerate overall imaging acquisition speed but increase these positioning errors.

We typically imaged GCaMP6s fluorescence in awake, behaving mice at an excitation wavelength of 920 nm with 60-95 mW power calibrated at the pia over 20-90 minute sessions, with little to no bleaching or change in baseline fluorescence. Suite2p reliably extracted neurons with >0.9 classifier probabilities, and nearly completely (i.e. to less than 0.1 pixels) removed x-y movement artifacts with uncorrected, raw principal components of up to ~1.3 pixels (i.e. up to ~15% of a 20 to 40 µm, or 4 to 8 pixels, diameter neuron at an imaging resolution of 0.2 µm/pixel, reduced to less than 0.5 µm or 1-2% of the diameter of a neuron after combined rigid and non-rigid motion correction).

To further control for z-movement artifacts, whose correction is intractable post-acquisition under normal scanning conditions, we also occasionally used GPU-enabled online fast-z motion correction (42/52 large FOV side mount sessions, and 9/39 large FOV dorsal mount sessions, used online z-correction) with ScanImage (Vidrio Technologies, LLC). This was only possible, in our hands, during multi-frame imaging when all frames were acquired at the same z-level. Quality of post-hoc motion-corrected dF/F traces and correlations with behavioral variables did not appear, by eye, to be different under these conditions (Fig. S3a vs b), as both sessions contained similar mixtures of neurons with walk-bout dependent, and walk-bout

independent, activity, for example. Because the z-motion correction did not seem to lead to large qualitative differences in neural activity and its relationship to behavioral primitives, both types of session are used interchangeably here.

In order to confirm neural region-of-interest (ROI) selection (i.e. each neuron as identified by a spatiotemporal footprint in Suite2p), increase scan speed, and correct our cortical layer 2/3 targeting for the effects of brain curvature, we also routinely imaged in random access mode with 4-6 small FOVs (660-1320 x 660 µm) positioned over visual (V1), somatosensory (S1), retrosplenial (RSpl), posterior parietal (PPC), and/or motor (MOs_pm, MOs_am, and/or MOs_al - ALM) cortices at independently controlled z-levels, with an imaging speed of 4-10 Hz and a resolution of 1.0 – 2.0 µm/pixel (see Protocol I; Fig. S2d, e). These sessions yielded hundreds of high quality neurons per FOV, at a density roughly twice that of the lower resolution sessions (Fig. S2c vs. d).

Uncontrolled image motion in behaving animals is a significant impediment to image quality (Pachitariu et al, 2016). We have taken significant care in our headpost design and fixation (see Figs. 1 a-b, S1 a-f, and Supplementary Methods and Materials) so that our preparations are highly stable, such that motion-corrected or “registered” movies showed very little to no residual movement in either large FOV, low-resolution, or multiple small FOV, high-resolution sessions (Suppl Movies S5, S7). Furthermore, movement-related activity across neighboring detected ROIs in large FOV, low-resolution sessions was correlated but non-identical upon detailed examination, and differed from dF/F Ca^{2+} fluorescence fluctuations detected in nearby non-ROI and blood-vessel regions of similar cross-sectional area (see Suppl Movies S5, S7). In theory, movement artifact signals should be relatively time-locked or nearly synchronous across cortical regions during vigorous walking or movement. However, we did not observe this in our preparations.

These findings, therefore, gave us confidence that neural activity changes detected during spontaneous movements of the mouse were due to real changes in Ca^{2+} fluorescence within each neuron, and not to movement of the brain causing the neuron to shift into or out of each Suite2p-defined ROI and leading to artifactual fluorescence transients (see Suppl Movies S5 and S7, which use blood vessels and inactive fluorophore/neuropil, respectively, as examples of the lack of significant movement-driven artifactual transients in GCaMP6s(-) structures in our preparations). In other studies performed using the same setup and methods as those applied here, we have demonstrated that non-activity dependent fluorescent markers, such as the expression of mCherry in cholinergic or noradrenergic axons, or autofluorescent blebs, exhibit no significant movement-related changes in fluorescence, indicating that

movement artifacts are minimal in our preparations (Collins et al, 2023). It should be noted that axonal diameters are on the order of ~1 µm or less - which is significantly smaller than the diameter of the neuronal cell bodies (~15-30 µm) imaged here. Axons, therefore, are a more stringent test of the stability of our imaging preparations.

Modulation of neural activity by spontaneous movements and arousal as observed with the dorsal mount preparation. Recent work has shown that sorting fluorescence based 2p GCaMP cortical neuronal activity with an algorithm called Rastermap is a powerful method to reveal clusters of cells that exhibit similar patterns of activity (Stringer et al, 2023, bioRxiv). An example of such sorting of our data for a dorsal mount session of 3096 neurons is shown in Figure 4a (top). This method for one-dimensional nonlinear embedding works by sorting and clustering of ROIs (i.e. neurons) by similarity of neural activity patterns across multiple timescales, using k-means clustering, sorting by asymmetric similarity as defined by peak cross-correlations at non-negative timelags, and upsampling of individual cluster activities in principal components space to allow sorting within clusters (Stringer et al, 2023, bioRxiv). Here, sorting with Rastermap revealed stark clustering of neurons with clearly distinguishable activity patterns that exhibited heterogeneous relationships to the behavioral primitive arousal/movement variables of walk speed, whisker motion energy, and pupil diameter (Fig. 4a top, b; Suppl Movie S4).

The activity generated by these thousands of neurons was clearly not random, with sub-groups of cells exhibiting similar activity patterns. Many neurons exhibited activity that appeared to be coupled to walking/whisking (cf. Fig. 4a top, b). It has previously been shown that the first principal component of cortical neuronal activity is highly correlated with movement (Stringer et al, 2019). Indeed, sorting our neurons based on degree of activity loading onto their first two principal components (i.e. neurons at the top of the sort have the largest amount of activity accounted for by the first principal component of activity across the entire population) revealed several distinct patterns of activity. The first principal component, PC1 (Fig. 4a, middle) demonstrated a clear relationship to walking/whisking, such that some neurons were strongly activated in association with movement (e.g. Fig. 4a, c1 in PC1, middle, and 4c, top), while other neurons appeared to be deactivated during movement bouts (e.g. Fig. 4a, c2 in PC1, middle, and 4c, bottom).

Closer examination revealed that some Rastermap and PC1 clusters were active primarily during walking (Fig. 4a, top, d1, Fig. 4a, middle, c1, and Fig. 4d1) or whisking, with either transient or sustained activity, while others were either active during low arousal periods

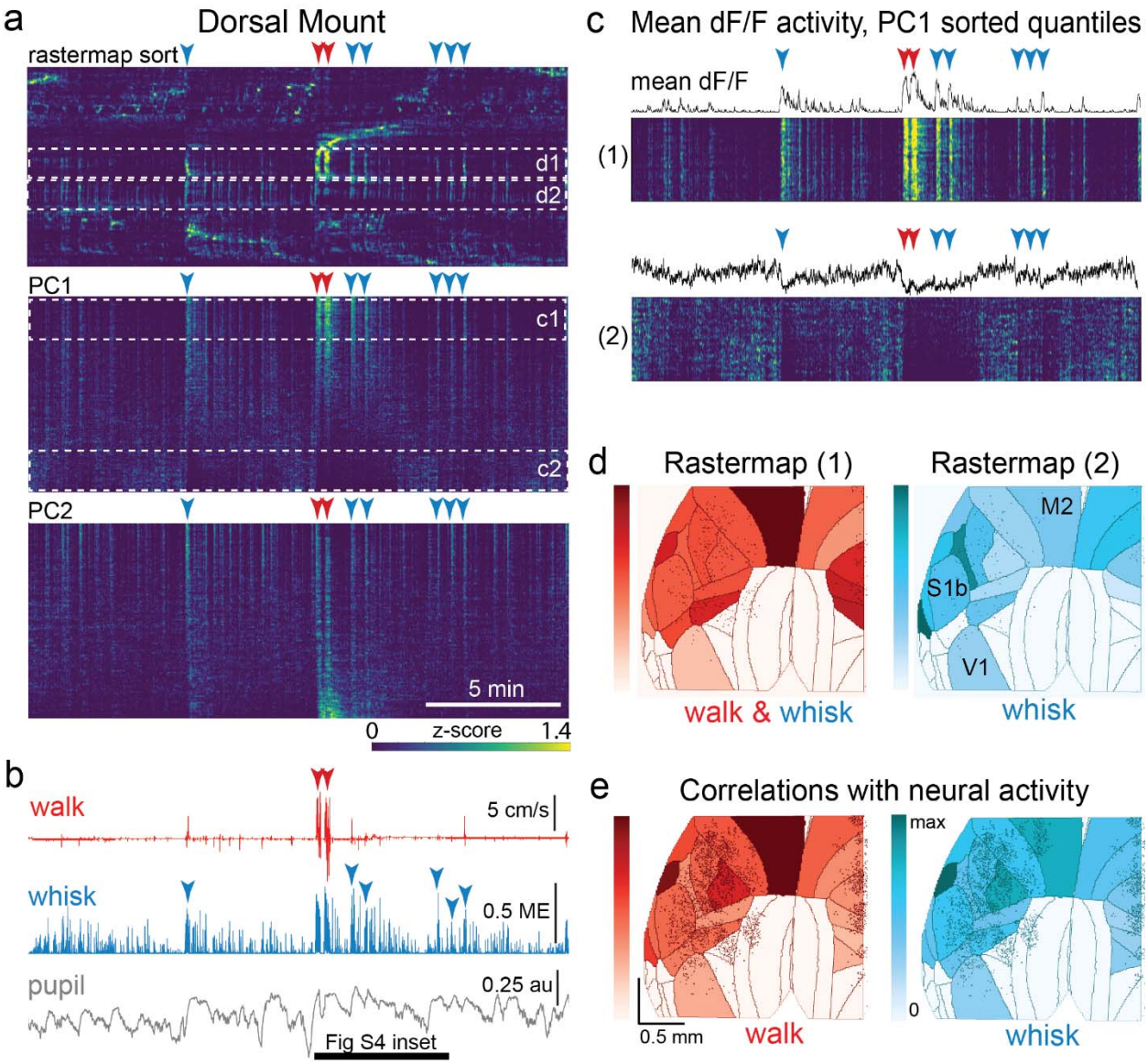

Figure 4. The relationship between spontaneous behavioral measures and neural activity across the dorsal cortex is globally heterogeneous. a) Rastermap (top), first principal component (PC1; middle), and second principal component (PC2, bottom) sorting of normalized, rasterized, neuropil subtracted dF/F neural activity for 3096 cells from a single 20 min duration example dorsal mount 2-photon imaging session. Each row in the display corresponds to a "superneuron", or average activity of 50 adjacent neurons in the Rastermap sort (Stringer et al, 2023, bioRxiv). Activity for each superneuron is separately z-scored, then displayed as low activity in blue, intermediate activity in green, and maximum activity level in yellow (standard viridis color look-up table). Red and blue arrowheads show alignment of walking and whisking bouts, respectively, to neural activity. White dashed box inserts indicate dual selections highlighted in panels c) and d). Scale bar shows color look-up map for each separately z-scored, individually displayed superneuron in the raster displays. **b)** Behavioral arousal primitives of walk speed, whisker motion energy, and pupil diameter shown temporally aligned to the rasterized neural activity traces in a), directly above. Horizontal black bar indicates time of expanded inset in Fig. S4. **c)** Expanded insets of top and bottom fifths of rasterized PC1 sorting from the middle segment of panel a), with mean dF/F activity traces shown above each. Red and blue arrowheads indicate the same walking and whisking bouts, respectively, as in a) and b). **d)** Normalized density of neurons in each CCF area belonging to two example Rastermap sorted groups (d1, left, red: MIN=0%, MAX=25%; d2, right, blue: MIN=0%, MAX=13%), with rasterized activities shown in corresponding labeled white dashed boxes in the top segment of panel a). The neurons present in selected Rastermap groups are shown as Suite2p footprint outlines. The type of behavioral arousal primitive activity typically concurrent with high neural activity for each Rastermap group is shown below each Rastermap group's CCF density map (i.e. walk and whisk for group

d1, left (red), and whisk for group d2, right (blue). **e)** Normalized mean correlations of neural activity and walk speed (left, red; mean: MIN=0.01, MAX=0.03) and whisker motion energy (right, blue; mean: MIN=0.02, MAX=0.08) per CCF area (i.e. average correlation of all neurons in each area) are shown for this example session. Mean walk speed correlations with neural activity (dF/F) were significantly more than zero ($p<0.001$, median $t(3095)=3.7$, single-sample t-test; python: scipy.stats.ttest_1samp) for 15 of the 19 CCF areas with at least 20 neurons present. The areas with mean correlations not significantly larger than zero were right VISam, right SSp_ll, left VIS_rl, and left AUDd. Mean whisker motion energy correlations with neural activity (dF/F) were significantly more than zero ($p<0.001$, median $t(3095)=4.4$, single-sample t-test) for 16 of the 19 CCF areas with at least 20 neurons present. The areas with mean correlations not significantly larger than zero were right VISam, left VISam, and left VISrl. PC = principal component, ME = motion energy, au = arbitrary units, M2 = secondary motor cortex, S1b = primary somatosensory barrel cortex, V1 = primary visual cortex.

(Fig. 4a, middle, c2) or appeared to display a combination of transient whisking-related activity (Fig. 4a, top, d2, and Fig 4d, right; see Fig. S4). In the illustrated example, alignment of marked increases and decreases in bulk population neural activity to eight brief walking bouts and corresponding whisking bouts and pupil dilations (red and blue arrowheads, respectively; Fig. 4a-c) was more clearly seen following ROI sorting by first and second principal components (Fig. 4a middle and bottom; PC1 and PC2, respectively, Fig. 4c) than in Rastermap sorting (Fig. 4a top).

Interestingly, separation of the top and bottom fifth of neurons in the example data set sorted by PC1 (Fig. 4a middle, c1 and c2, respectively; Fig. 4c top and bottom, respectively) showed large groups of neurons with clear positive (Fig. 4c (1), top) and negative (Fig. 4c (2), bottom) correlations with arousal/movement, as demonstrated by alignment of rasterized neural activity with walk and whisker transients, and local maxima in pupil diameter (see Fig. 4b). Specifically, neurons with high PC1 values (Fig. 4 c1) showed elevated activity during periods of increased arousal/movement, while neurons with low PC1 values (Fig. 4c2) showed depressed activity levels during these periods. Subsets of these neurons showed either elevated basal activity with suppression during periods of high activity corresponding to arousal and/or movement (Fig. 4c (2), top of Rastermap), or transient increased activity leading up to movement/arousal (Fig. 4c (2), bottom of Rastermap), followed by transient suppression (see Suppl Movie S4). Examination of individual “raw” dF/F traces confirmed that these differences were not due to artifacts of z-scoring or superneuron averaging (Stringer et al, 2023, bioRxiv) in the PC and Rastermap sorted displays (not shown).

Spatial distribution of neurons related to movement/arousal. Performing simple correlations between dF/F in each neuron with either whisker movements, walking, or pupil diameter revealed a broad range of values (Fig. S6b; whisk -0.4 to +0.6, pupil -0.6 to +0.6, and walk -0.2 to +0.4; same mouse as Fig.4, but a different session). The population distributions of these correlations were highly skewed (positive skew for whisk and walk, negative skew for pupil

diameter) with modes at small positive values (Fig. S6a, b). Plotting the spatial location of neurons by correlation value within the cortical map area revealed a broad distribution across the dorsal cortex of neurons whose activity was either strongly negatively or positively (or in between) correlated with walking, whisking, and/or pupil diameter, inclusively across every CCF subregion imaged (Fig. S6a-c).

At the local spatial scale (Fig. S6a, colored dots indicate individual neurons), neurons whose GCaMP6s activity was strongly positively correlated with movement or pupil diameter could often be found near neurons whose activity exhibited a wide variety of correlations, from strongly positive to strongly negative. Plotting the average correlation value for neurons within a CCF region revealed modest yet statistically significant (see Fig. S6c, legend) spatial heterogeneities (Figs. 4e, S6) that were evident both in each of two different single sessions (Figs. 4e; S6a, b) and across sessions (Fig. S6c, mean of 8 sessions from the same mouse).

To examine the spatial distribution of different Rastermap groups, we created heatmaps of the density of cells in each CCF area that belong to a particular Rastermap group (i.e. percent of cells in each region belonging to that Rastermap group; Fig. 4d; only areas with at least 20 cells were considered). A Rastermap group with increased activity in relationship to walking and whisking bouts (Rastermap cell group d1 in Fig. 4a, top, and S4a) had a CCF density distribution ranging from 0 to 25% and was concentrated in MOs, SSn, VISrl, and VISa (CCF neural density map; Fig. 4d, left). A second example Rastermap group, d2 (Figs. 4a, top, and S4c) exhibited increases in fluorescence in relationship to whisking, but not as strikingly during bouts of walking. This cell group had a CCF density distribution ranging from 0 to 13% and was concentrated in SSun, SSb, and AUDd (Fig. 4d, right).

Next, we compared the proportion of neurons within each CCF that exhibited large positive correlations with walking speed and whisker movement energy. In this example dorsal mount preparation these walk and whisk related neurons formed the largest relative proportion of cells in the CCF areas MOs, SSll, and SSn (Fig. 4e left and right, respectively). Mean correlations for each CCF area were generally statistically significantly greater than zero (see Fig. 4e, legend).

Modulation of neural activity by spontaneous movements and arousal as observed with the side mount preparation. To extend the potential diversity of neural activity patterns during spontaneous behavior, and to examine the generality of the observation of movement/arousal related neuronal activity to the lateral cortex, we monitored neuronal activity in spontaneously behaving mice with our lateral window preparation. We imaged ~85 sessions in the side mount

preparation (20-90 minutes each, including both large- and multi-FOV sessions) across 11 GCaMP6s mice, for a total of ~320,000 neurons (Fig. S2e; Suppl Movie S6). In an example imaging session (Fig. 5) of a mouse implanted with our custom 3D-printed titanium side mount headpost (Figs. 1e, S1f) and 10 mm radius bend cranial window (Figs. 1e, S1a, d, f; Suppl Movie S1), we were able to image 5678 neurons at 5 x 5 µm/pixel resolution over a combined 5.0 x 4.62 mm FOV (Fig. 5).

As in the dorsal mount (Fig. 4), Rastermap sorting of neurons across this single side mount example session (Fig. 5) was performed in order to further explore the broad range of relationships between movement/arousal measures and neural activity. As in the dorsal mount preparation, Rastermap sorting showed non-random patterns of neuronal activity that were often related to movement/arousal measures. Here, we identify and further characterize both a cluster with increased activity in relationship to bouts of walking/whisking (Fig. 5a, top, cell group d1; see Fig. S5; a, b), and another cell cluster whose activity appeared, by eye, to be more related to whisking than walking, *per se* (Fig. 5a, top, cell group d2; see Fig. S5; c, b). Interestingly, these two groups consisted of ~300 neurons each spread across the right lateral cortex, with some neurons nearby in the same CCF areas, and others in distant, functionally distinct regions (Fig. 5d left, right).

Sorting by the first principal component (PC1) of activity, as in our dorsal mount preparation (Fig. 4a, middle), revealed that a majority of imaged neurons exhibited activity related to bouts of walking/whisking (Fig. 5a, middle; walk bouts indicated by red arrowheads, whisk bouts indicated by blue arrowheads; see Suppl Movie S6). The top fifth of these movement-activated neurons showed activity nearly completely dominated by walk-related signals (Figs. 5a, middle, c1, and 5c, top, (1)), while the bottom fifth showed activity negatively correlated with walking (Figs. 5a, middle, c2, and 5c, bottom, (2)). Sorting by the second principal component (PC2), on the other hand, appeared to show increases in activity more closely aligned to the onset of whisking movements (Fig. 5a, bottom). Interestingly, some neurons in PC2 exhibited sustained inter-walk bout activation (Fig. 5a, bottom).

As in our dorsal mount example session (Fig. S6, a-c), we examined the overall spatial pattern of correlations in a different side mount example session with 3897 recorded neurons (Fig. S6, d-f). Here, correlations of individual neuron activity with whisker motion energy, pupil diameter, and walk speed ranged from -0.6 to +0.6 (Fig. S6d, e; whisk -0.4 to +0.6, pupil -0.4 to +0.4, and walk -0.2 to +0.5) with positively skewed distributions (except for pupil diameter) and modes at small positive values (Fig. S6d, e). Strong heterogeneity of these correlations, as in

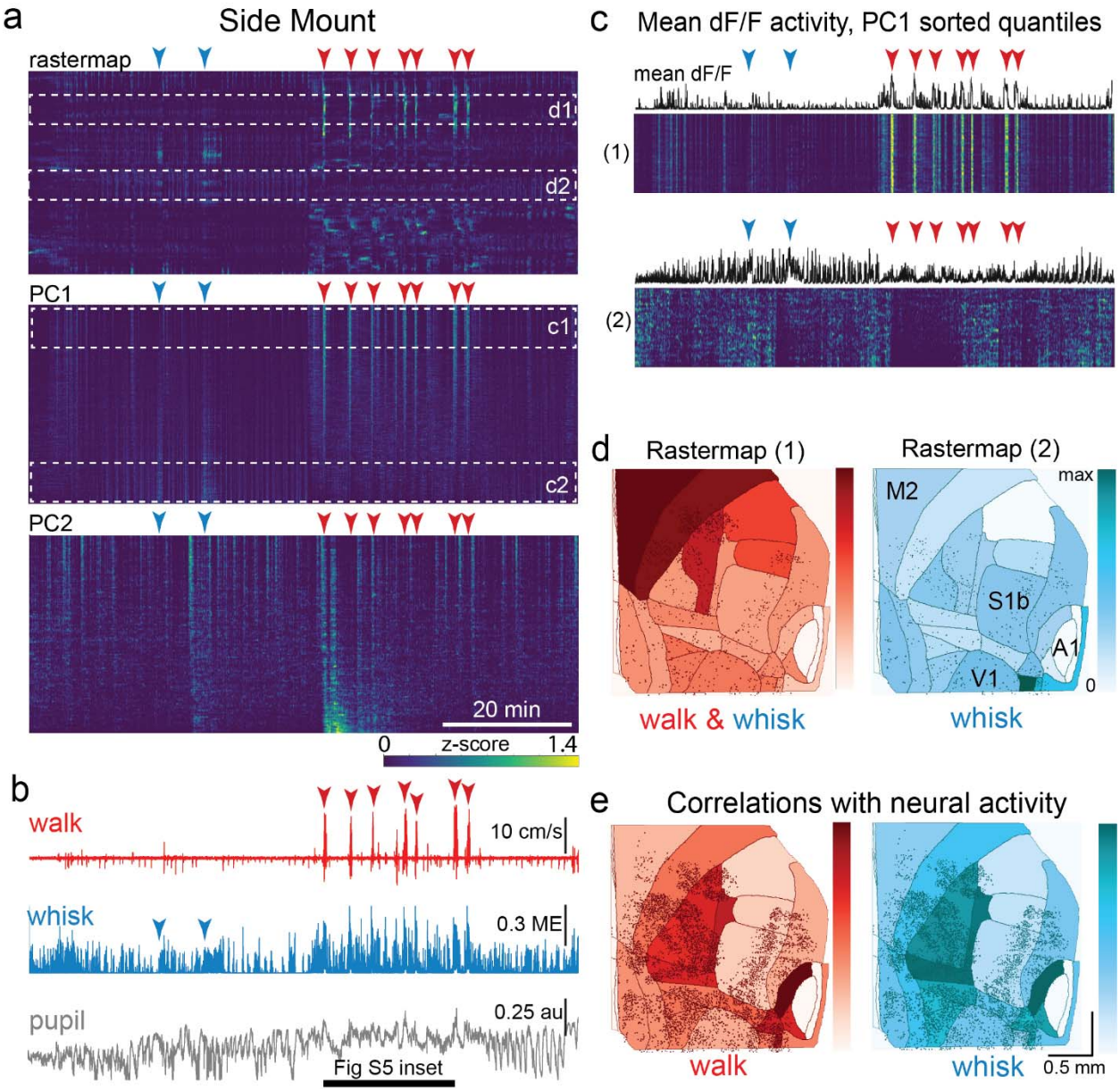

Figure 5. The relationship between spontaneous behavioral measures and neural activity across the lateral cortex is regionally patterned. a) Rastermap (top), first principal component (PC1; middle), and second principal component (PC2, bottom) sortings of normalized, rasterized, neuropil subtracted dF/F neural activity for 5678 cells from a single 90 min duration example side mount 2-photon imaging session. Each row in the display corresponds to a "superneuron", or average of 50 adjacent neurons in the Rastermap sort (Stringer et al, 2023, bioRxiv). Low activity (z-scored) in blue, intermediate activity in green, maximum activity level in yellow. Red and blue arrowheads show alignment of walking and whisking bouts, respectively, to neural activity. White dashed box inserts indicate selections highlighted in panels c) and d). Scale bar shows color look-up map for each separately z-scored, individually displayed superneuron in the raster displays. **b)** Behavioral arousal primitives of walk speed, whisker motion energy, and pupil diameter shown temporally aligned to the rasterized neural activity traces in a), directly above. **c)** Expanded insets of top and bottom fifths of rasterized PC1 sorting from the middle segment of panel a), with mean activity traces shown above each. Red and blue arrowheads indicate the same walking and whisking bouts, respectively, as in a) and b). Horizontal black bar indicates time of expanded inset shown in Fig. S5. **d)** Normalized density of neurons in each CCF area belonging to two example Rastermap sorted groups (d1, left, red: MIN=0%, MAX=34%; d2, right, blue: MIN=0%, MAX=32%), with rasterized activities shown in corresponding labeled white dashed boxes in the top segment of panel a). Only cells in selected Rastermap groups are shown. The type of behavioral arousal primitive (i.e. walk and whisk, left, in red; whisk, right, in blue) that was typically concurrent with

high neural activity is shown below each Rastermap group's CCF density map. **e)** Normalized mean correlations of neural activity and walk speed (left, red; mean: MIN=0.006, MAX=0.061; standard deviation: MIN=0.000, MAX=0.029) and whisker motion energy (right, blue; mean: MIN=0.009, MAX=0.079; standard deviation: MIN=0.000, MAX=0.043) per CCF area for this example session. Mean walk speed correlations with neural activity (dF/F) were significantly more than zero ($p<0.001$, median $t(5677)=4.4$, single-sample t-test; python: scipy.stats.ttest_1samp) for 20 of the 24 CCF areas with at least 20 neurons present. The areas with mean correlations not significantly larger than zero were left VISp, right SSpn, right AUDpo, and right TEa. Mean whisker motion energy correlations with neural activity (dF/F) were significantly more than zero ($p<0.001$, median $t(5677)=7.0$, single-sample t-test) for 21 of the 24 CCF areas with at least 20 neurons present. The areas with mean correlations not significantly larger than zero were right SSpn, right AUDpo, and right TEa. PC = principal component, ME = motion energy, au = arbitrary units, M2 = secondary motor cortex, S1b = primary somatosensory barrel cortex, V1 = primary visual cortex, A1 = primary auditory cortex.

the dorsal preparation example (Fig. S6, a-c), was evident at both local (Fig. S6d, e, single session) and regional (Fig. S6f, mean of 8 sessions from the same mouse) spatial scales.

Examining the regional variations in mean correlations between neural activity and walk speed (Fig S6f, bottom, in red), pupil diameter (Fig S6f, middle, in gray) and whisker motion energy (Fig 6Sf, top, in blue) showed that neurons with high correlations to arousal/movement tended to concentrate in M1 and somatosensory upper and lower limb (as one might expect for movement related activity) and dorsal auditory areas in this example session (dark red, gray, and dark blue, respectively), while neurons with lower correlations tended to concentrate in somatosensory barrel, nose, and mouth regions (light red, gray, and blue, respectively). As in the dorsal preparation (Fig. S6c), the mean correlations across 8 sessions from this example side mount mouse were in general statistically significantly greater than zero (see Fig. S6f, legend).

Behavioral video analysis and neurobehavioral alignment. From the above results, we can see that there are clear, but complex, relationships between behavior and neuronal activity across the dorsal and lateral cortex of the mouse. Principal component analysis of this activity suggests that one major factor is variations in behavioral state, as indicated by movement/arousal (see also Musall et al, 2019; Steinmetz et al, 2019; Stringer et al, 2019). Indeed, previous studies have demonstrated that the use of dozens or more of the top principal components of the motion energy of facial movement videos is beneficial in explaining variance in visual cortical activity (Stringer et al, 2019). However, beyond the first few principal components of facial movement, the precise behavioral meaning of the subsequent components are difficult to discern (Stringer et al, 2019).

Here we took a different approach by using a multi-step semi-automated analysis pipeline to identify higher order behavioral motifs and to examine the relationship of these motifs to the diversity of patterns of neural activity observed in the cortex.

To test the feasibility of this approach, we first created pose estimates consisting of (x,y) positions of a set of user-labeled body-parts and joints from our left, right, and posterior mouse videos using DeepLabCut (DLC; Mathis et al, 2018; see Suppl Movies S1). These pose estimate outputs were used for unsupervised machine learning extraction of behavioral motifs by an algorithm called "Behavioral segmentation of open field in DeepLabCut" ("B-SOiD"; Hsu and Yttri, 2021; see Suppl Movie S6). To identify high-level patterns of spontaneous head-fixed mouse behavior, we next performed manifold embedding ("UMAP") and cluster identification ("HDBSCAN") of DLC pose tracking outputs in BSOiD. A random-forests machine classifier was trained with these clusters to be able to predict, or identify, behavioral motifs in any test session with the same set or subset of labeled body part pose estimates (i.e. x-y positions for a certain number of body parts with the same names), either within or between mice.

In an example side mount session (same session as Fig. 5), we identified 16 BSOiD behavioral motifs that exhibited good coarse and fine alignment to the dynamics of behavioral arousal primitives (Fig. 6a), including walking, and to Rastermap groups from the same session. Here, BSOiD was trained on four sessions from the same mouse across three days, including the target session itself, with minimum cumulative motif times set to 1% of total video time or ~100 s, and then "tested" or used to label individual video frames from the target session. We found that, in some cases, training on individual sessions resulted in poor identification of rare and or brief behaviors, such as walking (which occurs rarely in some sessions; not shown).

Examination of four types of qualitatively identified behavioral epochs ("twitch", "whisk", "walk", and "pupil oscillate") at a finer temporal scale showed good alignment of BSOiD motifs with fine-scale features of spontaneous head-fixed mouse behavior at the level of movement and arousal primitives (Figs. 6a, b, Fig. S7). For example, balancing wheel movements during "twitch" seem to align well to flickering into BSOiD motifs 8 and 11, while flickering into motif 13 during "walk" appears to correspond well to individual periods of high walk speed during the extended walking bout (see Fig S7b). The qualitative labels that best describe these behaviors ("twitch", Fig. 6a-c; "whisk", Figs. 6a-c, S7; "walk", Figs. 6a-c, S7b; "pupil oscillate", Fig. 6a-c) did not appear to align with or be clearly explained by simple "first-order" interaction effects of movement and arousal primitives (i.e. whisk/walk/pupil x ON/OFF), despite the fact that the above-mentioned, closer, albeit preliminary, analysis of fine-scale behavioral features revealed precise and reliable alignments. This suggests that these methods will yield novel and important insights into the organization of neurobehavioral alignment across cortex, because other factors that are currently poorly understood appear to be contributing to the alignments.

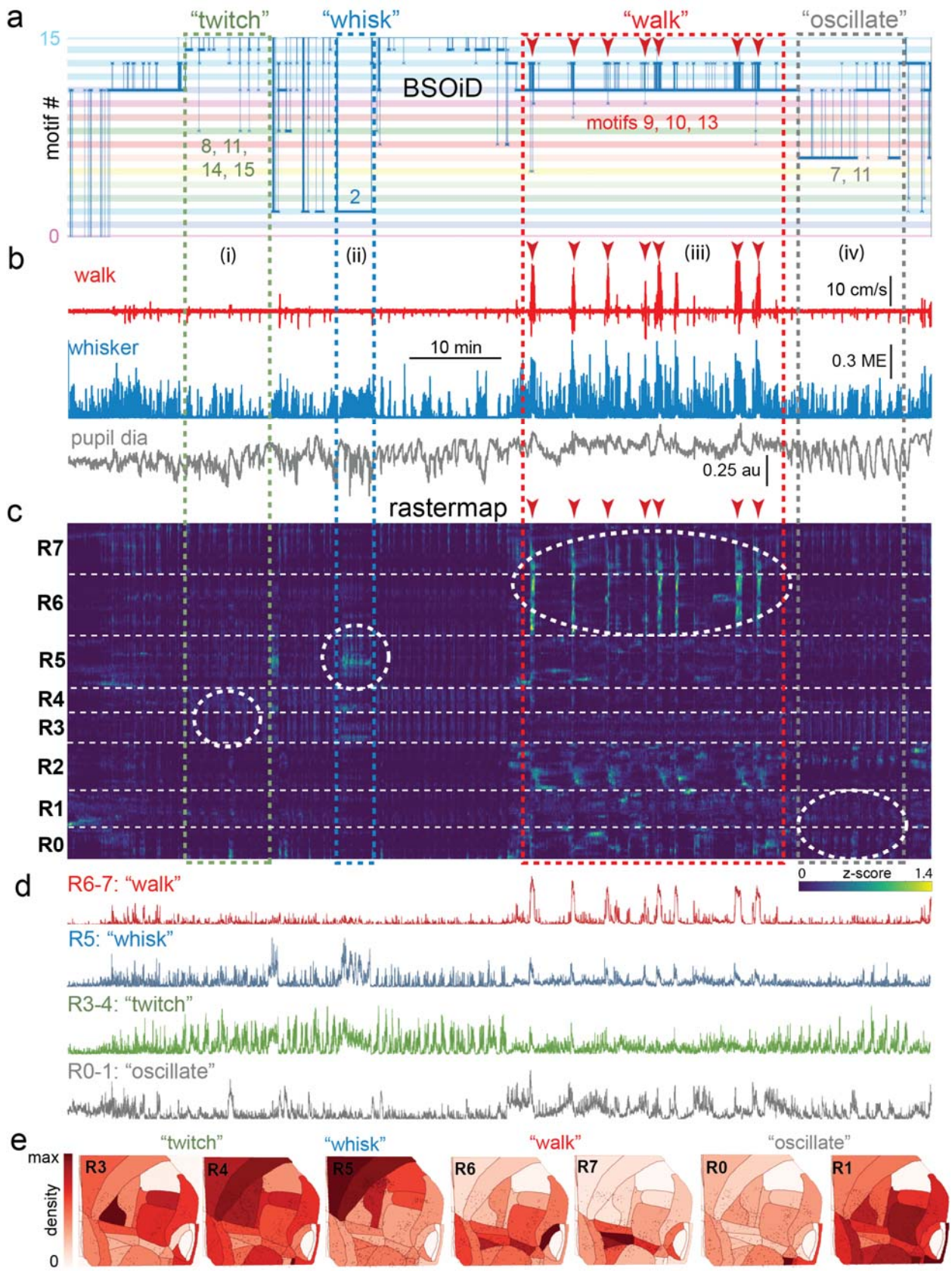
a
"twitch"
"whisk"
"walk"
"oscillate"
15
motif #
0
BSOiD
motifs 9, 10, 13
8, 11,
14, 15
2
7, 11
b
(i)
(ii)
(iii)
(iv)
walk
10 cm/s
whisker
10 min
0.3 ME
pupil dia
0.25 au
c
rastermap
R7
R6
R5
R4
R3
R2
R1
R0
0
z-score
1.4
d
R6-7: "walk"
R5: "whisk"
R3-4: "twitch"
R0-1: "oscillate"
e
max
density
0
"twitch"
"whisk"
"walk"
"oscillate"
R3
R4
R5
R6
R7
R0
R1

Figure 6. Alignment of activity sorted neural ensembles and spontaneous behavioral motifs reveals sparse, distributed encoding of arousal measures. a) Manual identification of high-level, qualitative behaviors (twitch, green dashed box; whisk, blue dashed box; walk, red dashed box; pupil oscillate, gray dashed box), aligned to sets of raw, unfiltered behavioral motifs (B-SOiD) extracted from pose-estimates (DeepLabCut) for a single example session. Numbered B-SOiD motifs (trained on a set of 4 sessions from the same mouse, motif # indicated by the position of the thick blue line relative to the y-axis at each time point) are color-coded (horizontal bars) across the entire 90-minute session. **b)** High level behaviors (dashed colored boxes, i-iv) and BSOiD motifs from a) are vertically temporally aligned to the behavioral primitive movement and arousal measures of walk speed, whisker motion energy, and pupil diameter. Expanded alignments of high level behaviors ii (whisk) and iii (walk) are shown in Fig S7a, b. Red arrowheads indicate temporal alignment positions for multiple walk bouts across BSOiD motifs (a), behavioral arousal primitives (b), and Rastermap sorted neural data (c). **c)** Rastermap sorted, z-scored, rasterized, neuropil subtracted dF/F neural activity (10th percentile baseline, rolling 30 s window) from the same side mount session, temporally aligned to the behavioral data directly above. Numbers at left (R0-R7) indicate Rastermap motif numbers, selected manually "by-eye" for this session, and horizontal dashed white lines on raster indicate separation between neighboring Rastermap motifs. Each row in the display corresponds to a "superneuron", or average of 50 adjacent neurons in the Rastermap sort (Stringer et al, 2023, bioRxiv). Colored, dashed boxes indicate alignment of high-level behaviors with neural activity epochs from the same example session. White dashed ovals indicate areas of Rastermap group-aligned active neural ensembles during periods of defined high-level behaviors. The 2-photon sampling rate for this session was 3.38 Hz. **d)** Normalized mean activity traces for all neurons in each neural ensemble indicated in (c) (white dashed ovals). Each trace also corresponds and is aligned to the indicated Rastermap groups. **e)** Color-coded neuron densities per CCF area corresponding to Rastermap groups shown in c), shown grouped by qualitative high-level behaviors, as indicated in a-b). Cross-indexing with neurobehavioral alignments (white dashed ovals) in c) allows for visualization of spatial distribution of neurons active during identified high-level behaviors (a) consisting of defined patterns of behavioral primitive movement and arousal measures (b). Corresponding maximum cell density percentages in each CCF area (white to red scale bar, top right) for each normalized Rastermap (0-7) heatmap color lookup table are 31.8, 24.0, 8.0, 11.4, 24.1, 38.4, 22.9, and 33.2%, respectively (left to right, top row followed by second row). This list of cell density percentages, therefore, corresponds to that of the CCF area filled with the darkest shade of red in each Rastermap group. White CCF areas in each Rastermap group are areas where no cells of that group were found, or where the total number of cells was less than 20 and therefore the density estimate was deemed unreliable and not reported. ME = motion energy, au = arbitrary units, CCF = common coordinate framework.

Alignment of the bulk activity of all neurons in each Rastermap group with BSOiD motif transitions for this example session (Fig. 6a-c, colored, vertically-oriented dashed boxes) showed differential alignment with the above-mentioned four high-level, manually identified qualitative behaviors ("twitch", "whisk", "walk", and "pupil oscillate"). The occurrences of these behaviors, then, were clearly aligned with both behavioral movement/arousal primitives (Fig. 6b) and the activity of putative, qualitatively segmented neural ensembles (Fig. 6c-d) that displayed varied spatial distributions across defined CCF areas of the side mount preparation (Fig. 6e). For example, "whisk", which corresponds to whisking in the absence of walking (Fig. 6a, b), with relatively elevated pupil diameter, aligns with strong activity in Rastermap group 5 (Fig. 6c-d, white dashed oval corresponding to Rastermap group #5, green/blue), which has a higher density of cells in M1, M2, and somatosensory upper limb, mouth, and nose CCF areas (Fig. 6e, R5).

We expect that, in general, the predictive contribution of high level behavioral motifs relative to movement and arousal primitives will be greater both in lateral areas, such as A1, in non-primary CCF areas further up the cortical hierarchy, such as secondary sensory cortices

and M2 (Wang et al, 2023; bioRxiv), and over intermediate and longer timescales of activity aligned with purposeful and/or goal oriented movements organized at higher levels of the behavioral hierarchy (Mimica et al, 2023).

Discussion

Overview and main findings. We developed novel headpost designs, surgical procedures, 3D-printed accessories, experimental protocols, and analysis pipelines that allowed us to routinely perform 2p imaging of up to 7500 neurons simultaneously across ~25 mm^2 of either parieto-occipital ("dorsal mount") or temporo-parietal ("side mount") cortex in awake, behaving mice, while simultaneously monitoring an array of global arousal state variables. Correlations between neural activity and arousal/movement were broadly, but heterogeneously, represented across both local and regional spatial scales. Preliminary analysis showed that this rich dataset could be used to segregate specific mappings between high-level, qualitatively identified behaviors (e.g. "twitch", "whisk", "walk", and "pupil oscillate"; Fig. 6a), each consisting of a persistent pattern of robustly identifiable, high-level behavioral motifs and low-level behavioral arousal, with movement primitives (i.e. walk, whisker, and pupil; Fig. 6b) and spatially localized neural activity clusters (i.e. "Rastermap" groups; Fig. 6c-e). Further analysis of such mappings will both allow for the fine dissection of patterned behaviors and neural activity, and for the development of a powerful framework for the sophisticated prediction of performance in mice engaged in a variety of multimodal sensory discrimination tasks (see Hulsey et al, 2023; bioRxiv, and Wang et al, 2023; bioRxiv).

Taken together, these findings show the strong potential of our novel methods for further elucidation of the principles of pan-cortical neurobehavioral alignment at the level of densely sampled individual neurons. For example, the spatial distributions of densities of neurons in each Rastermap (Stringer et al, 2023, bioRxiv) group across CCF areas were partially overlapping, yet strongly dissociable (see Figs. 4d, 5d, 6e). This suggested that dynamic, recurring patterns of arousal-dependent reorganization might act to enable activity mode switching between various distributed, cortical functional "communities." Each of these communities, then, could be specialized for different forms of activity related to various aspects of spontaneous behavior or task performance. This topic should be further explored with principled statistical techniques for recursive, top-down hierarchical community detection (Li et al, 2020).

In summary, we have shown here that it is feasible to monitor individual neuronal activity across broad expanses of the cerebral cortex and to perform neurobehavioral alignment of high resolution behavioral arousal state motifs and pan-cortical, activity-clustered neural ensembles in awake, behaving mice. Furthermore, our preliminary findings suggest a detailed alignment between both arousal/movement primitives and high-level behavioral motifs that appears to vary across broad regions of cortex (see Fig. 6). These findings extend those of earlier studies that showed widespread encoding of behavioral state during spontaneous behavior in visual cortex (Stringer et al, 2019; 2p imaging and Neuropixels recordings) and encoding of uninstructed movements during task performance both across dorsal cortex (Musall et al, 2019; 1p widefield imaging across dorsal cortex, 2p imaging in restricted FOVs in primary visual and secondary motor cortex) and brain-wide during task performance (at low volumetric sampling density; Steinmetz et al, 2019).

In particular, our findings are consistent with those of Stringer et al (2019) in that we observe widespread, strong correlations of cortical neural activity with the behavioral state variables of movement (facial and locomotor) and pupil diameter. However, we imaged larger areas of cortex simultaneously, including medial, anterior, and lateral cortical areas, and at a higher overall imaging frequency (~3 Hz vs. ~1 Hz). This may have contributed to our ability to observe enhanced heterogeneity across these non-primary cortical areas, including in neurons whose activities were strongly suppressed during periods of increased arousal/movement. Higher concentrations of such neurons were evident over lateral regions of cortex such as primary auditory and somatosensory nose and mouse regions, and in anterior regions including secondary motor cortex.

We also showed that high-level behaviors, manually identified and aligned to clusters of repeating motifs by UMAP embedding (BSOiD; Hsu and Yttri, 2021), aligned well with patterns of neural activity identified by Rastermap corresponding to neural ensembles with non-uniform spatial distributions (Fig. 6). This suggests that encoding of behavioral state across cortex may occur at multiple levels of the behavioral hierarchy, in addition to that of “low-level” behavioral arousal/movement primitives such as pupil diameter, whisker motion energy, and walking speed (see Mimica et al, 2023).

Our finding that neurons with no correlation to arousal/movement, or with a large negative correlation, are also widespread across cortex, suggests that multiple mechanisms or pathways may exist in parallel to distribute information about behavioral state across the brain. Interestingly, as evidenced by our discovery with Rastermap of neural ensembles exhibiting diverse response kinetics and polarities (Figs. 4a, top, 4d, 5a, top, 5d, and 6c-e), these

pathways may be activated with different temporal dynamics simultaneously across various brain areas (Shimaoka et al, 2018), thus making our approach for simultaneous recording of many areas one of the keys to potentially understanding the nature of their neuronal activity.

Additional experiments are warranted to uncover the mechanistic basis of such differences across large populations of cortical neurons associated with changes in activity level during periods of fluctuating arousal/engagement. For example, correlations between spiking and movement/arousal under 2p imaging, both as shown here and in previous studies (Musall et al, 2019; Stringer et al, 2019), and Neuropixels recordings (Steinmetz et al, 2019; Stringer et al, 2019), appear not to mirror those observed with widefield voltage imaging of membrane potential (Shimaoka et al, 2018; especially in VIS_lm and SS_b). This suggests the existence of a complex interaction between movement/arousal related changes in membrane potential and spiking that may vary across cortical areas.

In general, our preparations allow for the observation of diverse pan-cortical neural populations that likely communicate with each other in a transient or sustained manner during alternating periods of low and high movement/arousal and drive different patterns of global functional connectivity. Furthermore, the apparent variability of these neurobehavioral alignments suggests that interactions between cognitive and arousal/movement encodings may also occur during spontaneous behavior, in addition to during task performance (Musall et al, 2019), and that the degree of this interaction may depend on the cortical area.

Advantages and disadvantages of our approach. A main advantage of our approach is the flexibility afforded by the combination of our two surgical preparations (see Protocols II, III), in terms of the sheer number of cortical areas spanning distant regions that can be recorded simultaneously at single cell resolution under 2p imaging. The interoperability of our preparations across widefield 1p and Thorlabs mesoscope 2p imaging rigs (see Protocol IV) will allow for the development of an experimental pipeline that first surveys widespread activity at rapid timescales (widefield) and then investigates the activity of large subregions in a serial manner at intermediate timescales (Thorlabs mesoscope). In addition, we are currently working on modifying these preparations to allow combined 2p imaging and whole-cell electrophysiological recordings (i.e. using a Thorlabs Bergamo II microscope) and/or widefield imaging and high-density electrophysiological recordings (i.e. Neuropixels; Jun et al, 2017; Peters et al, 2021), so that a single preparation in each mouse can be used to acquire data across multiple spatiotemporal scales during spontaneous and task-engaged epochs.

A drawback of our current preparations is that they require use of a water-immersion objective with a relatively short working distance (~2.4-3.0 mm). This limits access for simultaneous electrophysiological recordings during mesoscale imaging, even though the same preparations could be used for simultaneous imaging and electrophysiological recordings on parallel acquisition rigs (i.e. widefield and/or standard 2p). Interestingly, recent advances in mesoscale FOV 2p microscopy (Diesel 2P; Yu et al, 2020) allow for the use of large working distance (~8 mm) air-immersion objectives. In tandem with our headpost, surgical, behavioral, and analytic protocols, this could enable simultaneous Neuropixels, whole-cell electrophysiology, and mesoscale single-cell resolution imaging experiments. Such an approach would allow integration of neural data from multiple brain depths and areas, as well as across multiple spatiotemporal scales.

Another limitation of our current methodologies as presented here is that our analyses were limited to neurons acquired at a single cortical depth, and at an imaging rate that did not take full advantage of the speed of the Ca^{2+} indicator that we used (i.e. GCaMP6s; we estimate that imaging at ~10 Hz would sample all available information with a rise time of ~200 ms and decay time of ~1.2 s; Chen et al, 2013). In the future we will expand our analyses to sessions, which we have already recorded, acquired at ~5-10 Hz over 4-6 FOVs (e.g. V1, A1, M2, RSpl, and SS_m/_n cortical areas; Fig. S2e), with resolutions ranging from 0.2 to 1 µm per pixel (Fig. S2d). Use of Dual Plane Imaging, a recently developed Thorlabs technology (developed together with the Allen Institute) that enables multiplexing of acquisition at more than one z-plane with no loss of imaging speed, with our preparations, in particular for use while pseudo-simultaneously imaging multiple small FOVs, would enable further in-depth examination of rapid fluctuations in neurobehavioral alignment structure across cortex in more than one cortical layer.

Future directions. Reliability of cortical area identification and alignment in our preparation is enhanced due to standardization of headposts and cranial windows, and the combined use of skull landmark and cortical vasculature alignment in a rigorous, semi-automated widefield multimodal sensory mapping protocol. This has allowed us to acquire an extensive, standardized neurobehavioral data set (~200 hours of spontaneous behavior with 30 Hz high resolution face and body video from three cameras, ~1,000,000 total neurons at ~3-10 Hz 2p acquisition rates imaged at ~1000 x 1000 pixels across ~250 sessions, total combined across both preparations and both FOV imaging configurations in 17 mice; see Fig. S2e).

In the future, we hope to expand our analyses of this extensive dataset to include identification of joint Rastermap/BSOiD (or keypoint-MOSEQ motif “syllable”; Weinreb et al,

2023; bioRxiv) transitions, and to perform further precision neurobehavioral alignment, using other methodologies such as hierarchical state-space models (Lindermann, S; https://github.com/lindermanlab/ssm), and factorial HMMs (Ghahramani and Jordan, 1997). These methods will allow increased accuracy over multiple behavioral timescales and the ability to predict transition or change points between different behavioral and/or neural activity states.

We will also examine changes in pan-cortical functional connectivity between communities of neurons with similar activity relationships associated with changes in behavioral state with statistical techniques such as Vector Autoregressive Union of Intersections (VA-UoI; Balasubramanian et al, 2020, Ruiz et al, 2020, Bouchard et al, 2022) that allow us to rigorously test for and remove putative but "false positive" connections based on coincidental activity correlations. Together, these analysis tools, when combined with the experimental techniques demonstrated here, will enable both identification of patterns of neurobehavioral alignment between arousal/movement and neural activity that are strongly predictive of high levels of behavioral performance, and the direct probing control of these patterns with targeted optogenetic manipulations.

References

1. Allen, W. E. *et al.* Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron* **94**, 891-907.e6 (2017).

2. Balasubramanian, M. *et al.* Scaling of Union of Intersections for Inference of Granger Causal Networks from Observational Data. in *2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* 264–273 (IEEE, 2020). doi:10.1109/IPDPS47924.2020.00036.

3. Bimbard, C. *et al.* Behavioral origin of sound-evoked activity in mouse visual cortex. *Nat Neurosci* **26**, 251–258 (2023).

4. Borisovska, M., McGinley, M. J., Bensen, A. & Westbrook, G. L. Loss of olfactory cell adhesion molecule reduces the synchrony of mitral cell activity in olfactory glomeruli: Olfactory

cell adhesion molecule and olfactory bulb synchrony. *The Journal of Physiology* **589**, 1927–1941 (2011).

5. Bouchard, K. E., Bujan, A. F., Roosta-Khorasani, F., Ubaru, S., Pabhat, Snijders, A. M., Mao, J-H, Chang, E. F., Mahoney, M. W., & Bhattacharyya, S. Union of Intersections (UoI) for Interpretable Data Driven Discovery and Prediction.. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA. USA.

6. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

7. Clancy, K. B., Orsolic, I. & Mrsic-Flogel, T. D. Locomotion-dependent remapping of distributed cortical networks. *Nat Neurosci* **22**, 778–786 (2019).

8. Coen, P., Sit, T. P. H., Wells, M. J., Carandini, M. & Harris, K. D. Mouse frontal cortex mediates additive multisensory decisions. *Neuron* S0896627323003811 (2023) doi:10.1016/j.neuron.2023.05.008.

9. Collins, L., Francis, J., Emanuel, B. & McCormick, D. A. Cholinergic and noradrenergic axonal activity contains a behavioral-state signal that is coordinated across the dorsal cortex. *eLife* **12**, e81826 (2023).

10. de Vries, S. E. J. *et al.* A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nat Neurosci* **23**, 138–151 (2020).

11. Esmaeili, V. *et al.* Rapid suppression and sustained activation of distinct cortical regions for a delayed sensory-triggered motor response. *Neuron* **109**, 2183-2201.e9 (2021).

12. Gallero-Salas, Y. *et al.* Sensory and Behavioral Components of Neocortical Signal Flow in Discrimination Tasks with Short-Term Memory. *Neuron* **109**, 135-148.e6 (2021).

13. Garrett, M. E., Nauhaus, I., Marshel, J. H. & Callaway, E. M. Topography and Areal Organization of Mouse Visual Cortex. *J. Neurosci.* **34**, 12587–12600 (2014).

14. Ghahramani, Z. & Jordan, M. I. Factorial Hidden Markov Models. *Machine Learning* **?**, 1-31 (1997).

15. Gong, Y. *et al.* High-speed recording of neural spikes in awake mice and flies with a fluorescent voltage sensor. *Science* **350**, 1361–1366 (2015).

16. Guo, Z. V. *et al.* Flow of Cortical Activity Underlying a Tactile Decision in Mice. *Neuron* **81**, 179–194 (2014).

17. Hattori, R., Danskin, B., Babic, Z., Mlynaryk, N. & Komiyama, T. Area-Specificity and Plasticity of History-Dependent Value Coding During Learning. *Cell* **177**, 1858-1872.e15 (2019).

18. Hsu, A. I. & Yttri, E. A. B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat Commun* **12**, 5188 (2021).

19. Hulsey, D., Zumwalt, K., Mazzucato, L., McCormick, D. A. & Jaramillo, S. *Decision-making dynamics are predicted by arousal and uninstructed movements.* http://biorxiv.org/lookup/doi/10.1101/2023.03.02.530651 (2023) doi:10.1101/2023.03.02.530651.

20. Jacobs, E. A. K., Steinmetz, N. A., Peters, A. J., Carandini, M. & Harris, K. D. Cortical State Fluctuations during Sensory Decision Making. *Current Biology* **30**, 4944-4955.e7 (2020).

21. Jun, J. J. *et al*. Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).

22. Kauvar, I. V. *et al.* Cortical Observation by Synchronous Multifocal Optical Sampling Reveals Widespread Population Encoding of Actions. *Neuron* **107**, 351-367.e19 (2020).

23. Kim, T. H. *et al.* Long-Term Optical Access to an Estimated One Million Neurons in the Live Mouse Cortex. *Cell Reports* **17**, 3385–3394 (2016).

24. Mathis, A. *et al.* DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21**, 1281–1289 (2018).

25. McGinley, M. J., David, S. V. & McCormick, D. A. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* **87**, 179–192 (2015).

26. Michaiel, A. M., Parker, P. R. L. & Niell, C. M. A Hallucinogenic Serotonin-2A Receptor Agonist Reduces Visual Response Gain and Alters Temporal Dynamics in Mouse V1. *Cell Reports* **26**, 3475-3483.e4 (2019).

27. Mimica, B. *et al.* Behavioral decomposition reveals rich encoding structure employed across neocortex in rats. *Nat Commun* **14**, 3947 (2023).

28. Morandell, K., Yin, A., Del Rio, R. T. & Schneider, D. M. *Movement-related modulation in mouse auditory cortex is widespread yet locally diverse*.

http://biorxiv.org/lookup/doi/10.1101/2023.07.03.547560 (2023) doi:10.1101/2023.07.03.547560.

29. Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience* **22**, 1677–1686 (2019).

30. Niell, C. M. & Stryker, M. P. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron* **65**, 472–479 (2010).

31. Pachitariu, M. *et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy*. http://biorxiv.org/lookup/doi/10.1101/061507 (2016) doi:10.1101/061507.

32. Peron, S. P., Freeman, J., Iyer, V., Guo, C. & Svoboda, K. A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. *Neuron* **86**, 783–799 (2015).

33. Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D. & Carandini, M. Striatal activity topographically reflects cortical activity. *Nature* **591**, 420–425 (2021).

34. Ruiz, T., Bhattacharyya, S., Balasubramanian, M., & Bouchard, K. Sparse and Low-bias Estimation of High Dimensional Vector Autoregressive Models. *Proceedings of Machine Learning Research* **120**, 1–10 (2020).

35. Salkoff, D. B., Zagha, E., McCarthy, E. & McCormick, D. A. Movement and Performance Explain Widespread Cortical Activity in a Visual Detection Task. *Cerebral Cortex* **30**, 421–437 (2020).

36. Shimaoka, D., Harris, K. D. & Carandini, M. Effects of Arousal on Mouse Sensory Cortex Depend on Modality. *Cell Reports* **22**, 3160–3167 (2018).

37. Shimaoka, D., Steinmetz, N. A., Harris, K. D. & Carandini, M. The impact of bilateral ongoing activity on evoked responses in mouse cortex. *eLife* **8**, e43533 (2019).

38. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *eLife* **5**, e14472 (2016).

39. Steinmetz, N. A. *et al.* Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science* **372**, eabf4588 (2021).

40. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019).

41. Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, eaav7893 (2019).

42. Stringer, C., Zhong, L., Syeda, A., Du, F. & Pachitariu, M. *Rastermap: a discovery method for neural population recordings*. http://biorxiv.org/lookup/doi/10.1101/2023.07.25.550571 (2023) doi:10.1101/2023.07.25.550571.

43. Sun, W. *et al. Learning produces a hippocampal cognitive map in the form of an orthogonalized state machine*. http://biorxiv.org/lookup/doi/10.1101/2023.08.03.551900 (2023) doi:10.1101/2023.08.03.551900.

44. Takita, T. *et al.* Chemistry of bleomycin. XXI. Metal-complex of bleomycin and its implication for the mechanism of bleomycin action. *J Antibiot (Tokyo)* **31**, 1073–1077 (1978).

45. Takita, T. *et al.* Chemistry of bleomycin. XXI. Metal-complex of bleomycin and its implication for the mechanism of bleomycin action. *J Antibiot (Tokyo)* **31**, 1073–1077 (1978).

46. The International Brain Laboratory *et al.* Standardized and reproducible measurement of decision-making in mice. *eLife* **10**, e63711 (2021).

47. Valley, M. T. *et al.* Separation of hemodynamic signals from GCaMP fluorescence measured with wide-field imaging. *Journal of Neurophysiology* **123**, 356–366 (2020).

48. Vesuna, S. *et al.* Deep posteromedial cortical rhythm in dissociation. *Nature* **586**, 87–94 (2020).

49. Wang, Q. *et al.* The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell* **181**, 936-953.e20 (2020).

50. Wang, Z. A. *et al. Not everything, not everywhere, not all at once: a study of brain-wide encoding of movement*. http://biorxiv.org/lookup/doi/10.1101/2023.06.08.544257 (2023) doi:10.1101/2023.06.08.544257.

51. Waters, J. Sources of widefield fluorescence from the brain. *eLife* **9**, e59841 (2020).

52. Weinreb, C. *et al. Keypoint-MoSeq: parsing behavior by linking point tracking to pose dynamics*. http://biorxiv.org/lookup/doi/10.1101/2023.03.16.532307 (2023) doi:10.1101/2023.03.16.532307.

53. Wu, Z. *et al.* Context-Dependent Decision Making in a Premotor Circuit. *Neuron* **106**, 316-328.e6 (2020).

54. Xiong, Q., Znamenskiy, P. & Zador, A. M. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* **521**, 348–351 (2015).

55. Yoshida, E. *et al.* In vivo wide-field calcium imaging of mouse thalamocortical synapses with an 8 K ultra-high-definition camera. *Sci Rep* **8**, 8324 (2018).

56. Yu, C.-H., Stirman, J. N., Yu, Y., Hira, R. & Smith, S. L. *Diesel2p mesoscope with dual independent scan engines for flexible capture of dynamics in distributed neural circuitry*. http://biorxiv.org/lookup/doi/10.1101/2020.09.20.305508 (2020) doi:10.1101/2020.09.20.305508.

57. Zhou, M. *et al.* Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat Neurosci* **17**, 841–850 (2014).

58. Zhuang, J. *et al.* An extended retinotopic map of mouse cortex. *eLife* **6**, e18372 (2017).

Acknowledgements.

We thank Luca Mazzucato for helpful comments on the manuscript and Paul Steffan, Lawrence Scatena, and Daniel Hulsey for technical assistance. We thank Jack Waters for contributing CCF outlines and masks for the side mount rotated view cortical map. We thank Kazi Rafizullah, John Boosinger, and Eowyn Boosinger for helping conceive, design, and build the mesoscope resonant scanner noise reduction shield. We also thank Elliott Abe, David Wyrick, Shreya Saxena, Matthew Kaufman, Alexander Hsu, Caleb Weinreb, Jens Tillmann, and Carsen Stringer for assistance with coding and setting up preliminary data analysis for both neural and behavioral data, and the teams at Vidrio ScanImage and Thorlabs for technical assistance with software and hardware related to imaging acquisition, respectively. This work was supported by NIH grants R35NS097287 (DAM) and R01NS118461 (DAM).

Author contributions.

E.D.V and D.A.M. conceived and designed the study. E.D.V. performed the hardware and protocol design, surgeries, data acquisition, and data analysis. E.D.V and D.A.M. wrote the paper.

Supplementary Files.

All supplementary analysis code, 3D-printable files, and other design files are publicly available at: https://github.com/vickerse1/mesoscope_spontaneous. Please cite this work if you use them.

Declaration of interests.

The authors declare no competing interests.