Running modulates primate and rodent visual cortex via common mechanism but quantitatively distinct implementation

John P. Liska[1], Declan P. Rowley[1], Trevor T. K. Nguyen[1], Jens-Oliver Muthmann[1], Daniel A. Butts[2], Jacob L. Yates[2,3,+,*], and Alexander C. Huk[1,+]

[1]Center for Perceptual Systems, Institute for Neuroscience, Department of Neuroscience, Department of Psychology, The University of Texas at Austin, Austin TX, 78712, USA
[2]Department of Biology and Program in Neuroscience and Cognitive Science, University of Maryland, College Park MD, 20742, USA
[3]Herbert Wertheim School of Optometry and Vision Science, University of California, Berkeley, CA, 94704, USA
[*]Corresponding author: yates@berkeley.edu
[+]these authors contributed equally to this work

When mice actively locomote, visual signals in their primary visual cortex (V1) are strongly modulated[1]. This observation has fundamentally altered conceptions of a brain region previously assumed to be a passive image processor, and extensive work has followed to dissect the sources, recipients, and functional consequences of running-correlated modulation[2–13]. However, it remains unclear whether visual processing in primates might similarly change during active locomotion. We therefore measured V1 activity in a nonhuman primate, the common marmoset (*Callithrix jacchus*), while they alternated between running and stationary. In contrast to the large increases in mouse V1 during running, conventional metrics of response in marmoset V1 were slightly but reliably decreased during running. However, by leveraging large-scale recordings, analysis of the latent variables driving population activity revealed a common mechanism in both species: trial-to-trial fluctuations of shared gain modulations were present across V1 in mice and marmosets. These gain modulations were larger in mice and were often positively correlated with running; they were smaller and more likely to be weakly negatively correlated with running in marmosets. Thus, population-scale gain fluctuations of V1 reflect a common principle of mammalian visual cortical function, but there are important quantitative differences in their magnitudes and correlations with behavior that yield distinct consequences for the relation between vision and action in primates versus rodents.

Primary sensory areas, such as primary visual cortex (V1), are by definition the first cortical stages to receive input from the corresponding sensory receptors, and are therefore often thought of as solely driven by sensory inputs. Neurons in V1 of virtually all mammals are known to be selective for image orientation within a restricted region of the visual field, a presumably critical early step of image processing that continues throughout a hierarchy of visual brain areas[14,15]. Mice, one of the major mammalian model systems in neuroscience, have neurons in V1 that are driven strongly by visual input and exhibit canonical orientation selectivity[16]. Recent work has, however, demonstrated that mouse V1 responses are also powerfully affected by whether or not the mouse is stationary or is running, while head-fixed on a treadmill or ball[1]. This has challenged the notion of this earliest cortical stage of visual processing as being purely or even primarily sensory in nature.

However, these observations have all been made in rodents, and have not been addressed at the single-cell level in primates. Although rodents certainly rely on vision for important behaviors[17], primates are more fundamentally visual organisms, with exquisite acuity and specialized functional characteristics such as foveas and corresponding high-resolution representations of the central visual field in V1[18]. We therefore tested the generality of running-based modulations in the V1 of the marmoset, a highly visual new world primate. We recorded from ensembles of neurons while marmosets on treadmills viewed stimuli matched to mouse V1 experiments, and used the same analysis pipeline to compare primate and rodent datasets at the level of individual sessions and units. Marmosets were head-fixed, placed on a wheel-based treadmill, and alternated between running and not running while we presented various visual stimuli designed to assess the properties and responsiveness of V1 neurons (Figure 1a,b). We recorded from foveal and parafoveal neurons in 2 marmosets (using chronically-implanted N-Form 3D electrode arrays), and were also able to simultaneously record from both foveal and peripheral V1 in one marmoset (using Neuropixels 1.0 probes). To support exacting comparison to rodent V1, we accessed a publicly-available mouse dataset that used matching stimuli in a treadmill paradigm[19]. This let us analyze our marmoset data and this mouse dataset identically, and thus to perform direct statistical comparisons between the species.

First, we mapped the receptive fields of marmoset V1 neurons using reverse-correlation techniques adapted to free-viewing[20] while we measured gaze using a video-based eyetracker (Fig 1c). In V1 of both marmosets, we found receptive fields within the

central few degrees of vision, with sizes expected at those eccentricities (1-5 deg, Fig 1f, blue and green; these can be compared to those in mouse, Fig 1e). Also as expected for primary visual cortex, marmoset V1 (in both well-isolated single units and well-tuned multi-unit clusters) responded robustly to oriented gratings and exhibited orientation- (and sometimes direction-) selectivity[21,22], similar to that in the mouse V1 dataset (Figure 1g,h). Orientation tuning spanned a range from weak to strong tuning, with many units exhibiting strong and conventional tuning curves (Fig 1i,j).

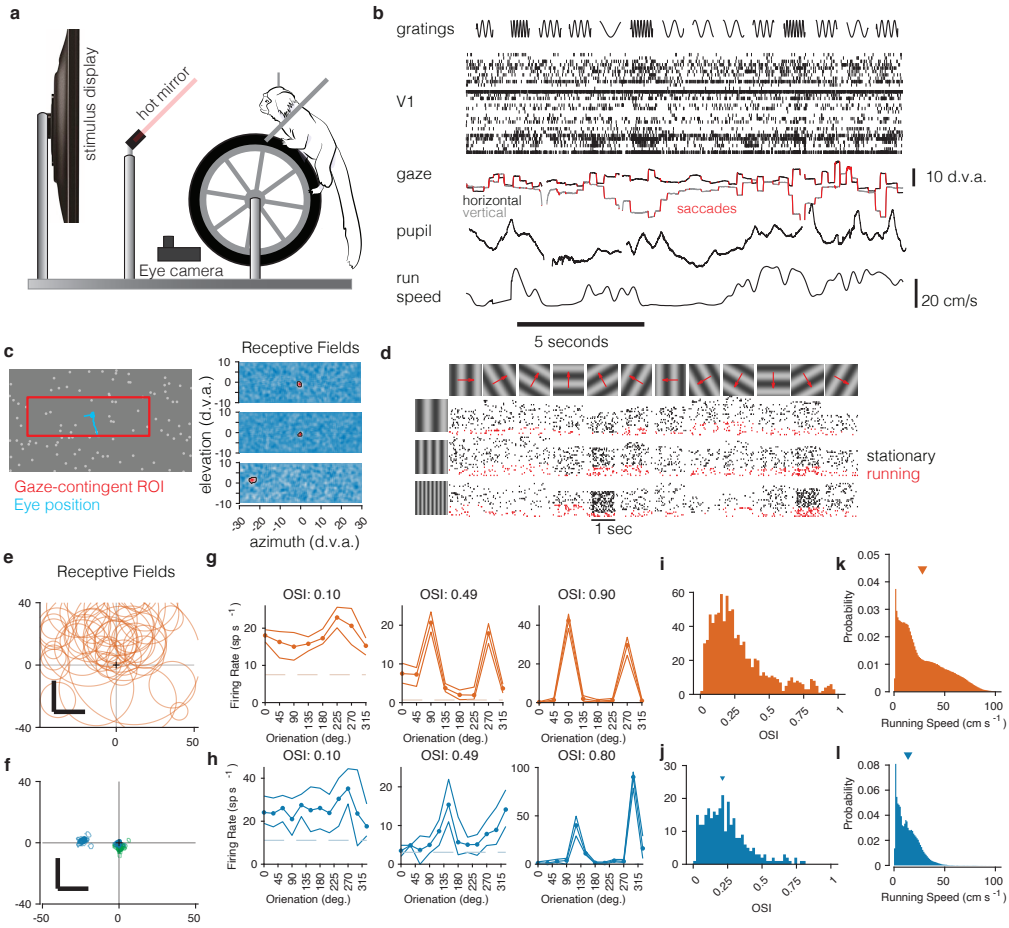

Figure 1. Recording from marmoset V1 during active locomotion. a) Apparatus for recording from marmoset V1 while presenting visual stimuli on a high-resolution display, monitoring gaze using an eye tracker, on a circular treadmill that allowed the marmoset to run or not run. Marmosets were acclimated and motivated to run with fluid rewards yoked to traveling a criterion distance. b) Schematic example of variables of interest. Visual stimuli were presented (top row). Rasters show activity from a V1 array (second row). Gaze was monitored (3rd row, x and y timeseries plotted in black and grey), saccades were detected (red), and pupil size was also measured (4th row). Running speed was measured using a rotary encoder attached to the treadmill (5th row). c) Before the main experiments, receptive fields were mapped using sparse noise[20]. The array of pseudocolor images shows 3 examples of V1 receptive fields (2 foveal and 1 peripheral neuron). d) Main experiment involved presenting full-field sinusoidal gratings that drifted in one of 12 directions (top row), at a variety of spatial frequencies (vertical axis at left). Rasters show example V1 activity during stimulus presentations when running (red) or stationary (black). e) Summary of receptive field (RF) locations in the mouse dataset (orange, top), and f) our data from marmosets (blue and green, bottom). In both marmosets, we recorded from a portion of V1 accessible at the dorsal surface of the brain using chronically implanted arrays, which yielded neurons with foveal receptive fields (green RFs). We also recorded from marmoset 1 using Neuropixels arrays, which allowed us to access both peripheral and foveal V1 (blue RFs; peripheral units are analyzed later/separately, see text). g) Examples of mouse V1 orientation tuning curves, for cells with weak, moderate, and strong orientation tuning. h) Same, for marmoset V1. i,j) Histograms of orientation-selectivity indices (OSIs) for mice (i) and marmosets (j). Marmoset OSIs, likely lower than previously reported because we used full-field stimuli not optimized to the spatial frequency tuning of each neuron. Regardless, the marmoset V1 neurons had strong visual responses and qualitatively conventional tuning. k,l) Running speeds in mice (k) and marmosets (l).

As a first test for effects of running on V1 activity, we assessed whether running speed was correlated with aggregate V1 activity by comparing the time series of these variables throughout each session. In the mouse, such modulations are easily visually evident when inspecting the time series of neural activity and running: when the mouse runs, V1 spiking often increases substantially. Fig 2a,b show example sessions with the maximal and median amounts of correlation between the time series of running speed and a generic low-dimensional representation of the population activity (the first principal component [PC] of the simultaneously-recorded V1 trial spike counts). This correlation could be seen when running / not running alternated on slow (Fig 2a) or fast (Fig 2b) time scales.

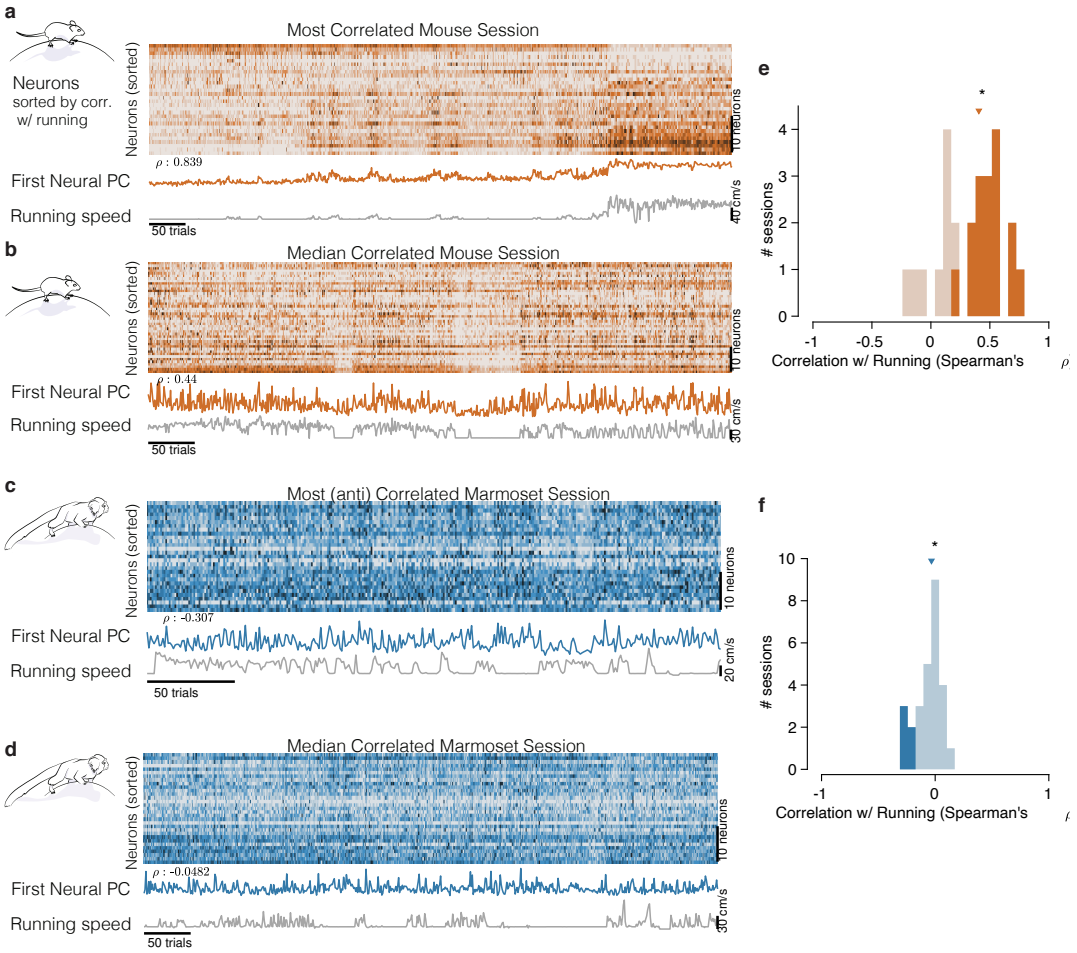

Figure 2. Mice and marmosets exhibit different correlations between V1 activity and running speed, evidenced on a session-by-session basis. a) Mice show visually compelling correlations between V1 trial spike counts and running speed. Example session with the highest correlation between running and V1 activity. Raster at top shows spiking activity of all mouse V1 neurons recorded. Population activity is summarized below the raster as the first principal component of the V1 array activity ("First Neural PC", orange trace); running speed is plotted underneath it on the same time axis (grey trace). Clearly, the two curves are highly similar. b) Same, for an example mouse session chosen to have the median correlation between running and V1 activity. In this example, the modulations of running speed and neural activity rise and fall together on a faster time scale than in the example in a). e,f) Marmosets show smaller, and typically negative, correlations between V1 spiking activity and running speed. Format same as the mouse data in (a,b), with example sessions chosen to show the maximal and the median correlations between V1 activity and running speed. The (anti-correlated) similarity between V1 activity (First Neural PC) and Running Speed curves is harder to discern in the marmoset. e,f) Correlations between V1 activity and running in the mouse (e) had a median > 0 (median=0.407, p=9.04×10^{-5}, stat=308, n=25, Mann-Whitney U Test), and many individual sessions had significant correlations with running (filled bars), and all such significant sessions had positive correlations. In the marmosets (f), the distributions of correlations were slightly but reliably negative (median=-0.033, p=0.034, stat=101, n=27, Mann-Whitney U Test), and all significantly modulated individual sessions exhibited negative correlations (5/27).

Relations between running and the activity of marmoset V1 neurons representing central vision were considerably smaller and less compelling. A starkly different impression comes from visual inspection of sessions with maximal and median relations between running and V1 activity (Fig 2c,d show example sessions with the maximal and median correlations). In these examples, V1 activity did not track running speed as clearly, although the activity did tend to increase when the monkey stopped running, explaining the modest negative correlations.

We then quantified these relations across all experiments on a session-by-session basis, in both species. For mice, this confirmed a strong positive correlation (Fig 2e; median=0.407, n=25, p=9.04×10^{-5}, stat=308, Mann-Whitney U Test). For marmosets, the distribution of correlations between V1 activity and running was subtly but reliably negative (Fig 2f, median=-0.033, p=0.034, stat=101, n=27, Mann-Whitney U Test). The correlation with running was significantly different between the two species (p=6.93×10^{-7}, stat=934, Mann-Whitney U Test). This session-level analysis confirmed that running modulations in mice are large and mostly reflect increases in response. Running modulations in marmoset foveal V1 are slightly suppressive.

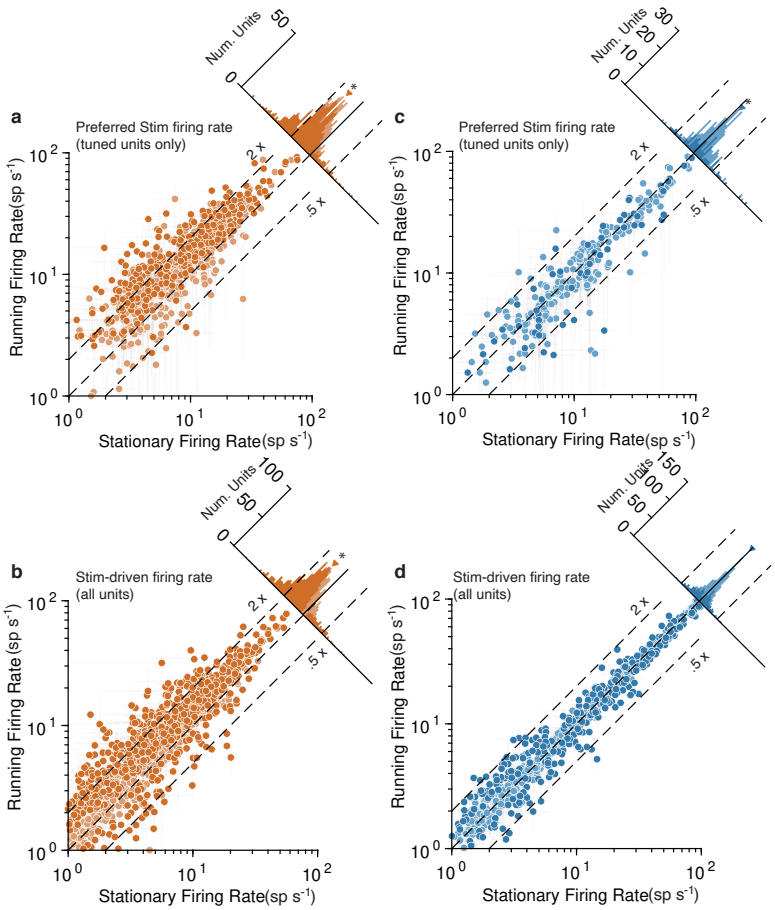

Figure 3. Running strongly increases mouse V1 activity and subtly decreases marmoset V1 activity, evidenced at the level of individual units. The mouse data points are plotted in orange and the marmoset data in blue. a) Scatterplot (log-log) shows firing rate to preferred stimulus for tuned units (OSI > 0.2), during running (y-axis) and stationary (x-axis). Histogram summarizes the projections onto the line of unity, and shows a clear shift indicating increases in response during running (geometric mean ratio [running/stationary] = 1.523 [1.469, 1.579], n=743). Dark-shaded symbols indicate individually-significant units. Dashed lines indicate doubling (2X) and halving (0.5x) of response. b) Same format, but now showing the response aggregated over all stimuli, for all units (geometric mean ratio [running/stationary] = 1.402 [1.365, 1.440], n=1168). A similar pattern reflecting primarily large increases is evident. c,d) V1 units in marmoset show a very different pattern. Responses of tuned units to preferred stimuli (c) cluster more closely to the line of unity, with a small but significant shift indicating a subtle decrease in response (geometric mean ratio [running/stationary] = 0.899 [0.851, 0.949], n=228). Responses to all stimuli for all units (d) show even less running-related modulation (geometric mean ratio [running/stationary] = 1.011 [0.995, 1.027], n=786).

To perform additional quantitative tests at the level of individual V1 units, we took the spike rate responses to drifting gratings of variable orientations, and divvied them up based on whether or not the marmoset was running (Figure 3). This analysis confirmed, in mouse, a tendency for large response increases during running to both the preferred (Fig 3a, geometric mean ratio [running/stationary] = 1.523 [1.469, 1.579], n=743 tuned units) and to all visual stimuli (Fig 3b, geometric mean ratio [running/stationary] = 1.402 [1.365, 1.440], n=1168). Many individual units had significant running modulations, and were more often increases rather than decreases (803/1168 [68.75%] increased firing rate, and 115/1168 [9.85%] decreased firing rate, bootstrapped t-test). In marmoset V1, there was again a modest decrease evident in the response to the preferred stimulus (Fig 3c; geometric mean ratio [running/stationary] = 0.899 [0.851, 0.949], n=228 tuned units). This suppression was less evident in responses aggregated across all stimuli (Fig 3d, geometric mean ratio [running/stationary] = 1.011 [0.995, 1.027], n=786). The number of significantly modulated units was relatively small, and was more balanced between decreases and increases in response (172/786 [21.88%] increased firing rate, and 161/786 [20.48%] decreased firing rate, bootstrapped t-test). Because we performed quantitative comparisons on subsets of the data for which the stimuli were nearly identical across species, and used the same data analysis programs to calculate response metrics, these analyses solidly confirm a substantial difference between the form of running modulations of V1 activity in mouse versus marmoset (log ratio of running:stationary was significantly different between mouse and marmoset for all units: $p=6.62 \times 10^{-99}$, stat=1399874, Mann-Whitney U Test, and tuned units: $p=4.69 \times 10^{-59=7}$, stat=4030135, Mann-Whitney U Test). Thus, the aggregate impacts of running on V1 responses, expected to have downstream impacts on perception and action, were much larger in mice and of opposite sign in marmoset.

Given these apparently categorical differences between the two species at the levels of both experimental sessions and individual units, we then employed more powerful neuronal population analyses to gain mechanistic insight into the nature of running-based modulations. A model incorporating trial-to-trial variation in shared gain modulations across V1 (Figure 4a) improved descriptions of the population data over simpler models that only took the stimulus (and slow drifts in baseline firing rate) into account for all sessions (figure 4d)[23] (figure 4b; marmoset $p=1.52 \times 10^{-82}$, stat=27174, n=754, Wilcoxon signed rank test; mouse $p=4.64 \times 10-181$, stat=25966, n=1257). This was true in both species, bolstering the emerging notion that population-level gain modulations are a general principle of mammalian V1 function[23–25,25–27]. This shared gain term, estimated only from the simultaneously-recorded neural activity (i.e., with no appeal to behavioral state), modulated more strongly in mice compared to marmosets (figure 4e, std. dev. in mouse = 2.170 [2.106, 2.245], marmoset = 1.188 [1.072, 1.274], $p<1 \times 10-9$, stat=1013202, Mann-Whitney U Test). Furthermore, in the mouse, shared gain was higher during running than stationary stimulus presentations (mean difference 0.970 [0.761, 1.225], p 0, stat 8.017, t test)– demonstrating that a substantial portion of modulations of mouse V1 can be explained by a shared gain term that increases with running. In marmoset, however, shared gain was slightly but reliably lower when running (mean difference = -0.125 [-0.203, -0.059], p=0.002, stat=-3.360, t test) a quantitatively very different relation to running than in mouse ($p = 8.77 \times 10^{-9}$, stat=6.615, 2 sample t test).

Although our marmoset dataset focused on V1 neurons representing the central portion of the visual field, we were able to record simultaneously from neurons with peripheral and foveal receptive fields by advancing a Neuropixel probe into both the superficial portion of V1 (foveal) and the calcarine sulcus (peripheral), resulting in simultaneous recordings of 110 and 147 (stimulus-driven) units representing the central and peripheral portions of the visual field, respectively. Analyzing neurons with peripheral receptive fields separately revealed a difference in running modulations between these retinotopically-distinct portions of V1: peripheral neurons had slightly higher stimulus-driven responses during running (aggregating over all stimuli, geometric mean ratio [running/stationary] = 1.129 [1.068, 1.194], n=147; difference was significant, p=2.100e-03, stat=12376, Mann-Whitney U Test), and the two sessions in which we were able to perform these measurements had higher positive correlations than any sessions in our entire foveal V1 dataset (assessed by correlating running speed either with the First Neural PC or with a shared gain term). Although the foveal representation in V1, which is relatively specific to primates, is slightly suppressed by running, it appears that quantitative differences exist in the peripheral representation. Regardless, our initial measurements suggest that the correlations with running in the periphery are still small relative to those in mouse V1 (median spike rate modulation by running significantly different between mouse and marmoset calcarine recordings: $p=7.639 \times 10^{-11}$, stat=7967825, Mann-Whitney U Test).

Our results support mechanistic insights into cross-species comparisons of V1 activity and running. The quantitative difference in the magnitudes and signs of V1 gain modulations are consistent with known differences in neuromodulatory inputs related to arousal in rodent and primate V1[28,29]. In primates, the locations of ACh receptors allow cholinergic inputs to increase the activity of the majority of GABAergic neurons and hence suppress net activity via inhibition[30,31], but pharmacologically and anatomically distinct cholinergic influences in rodent likely exert more complex effects on net activity, including disinhibition which can increase net activity[8,10,32]. Our population-level analyses also forge connections to indirect and aggregate measures of neural activity made in humans under related conditions[33–35], as well as the typically small modulations seen in primate visual cortices elicited by carefully-controlled attentional tasks, which are more clear when ensemble-scale modulations are considered[36–38].

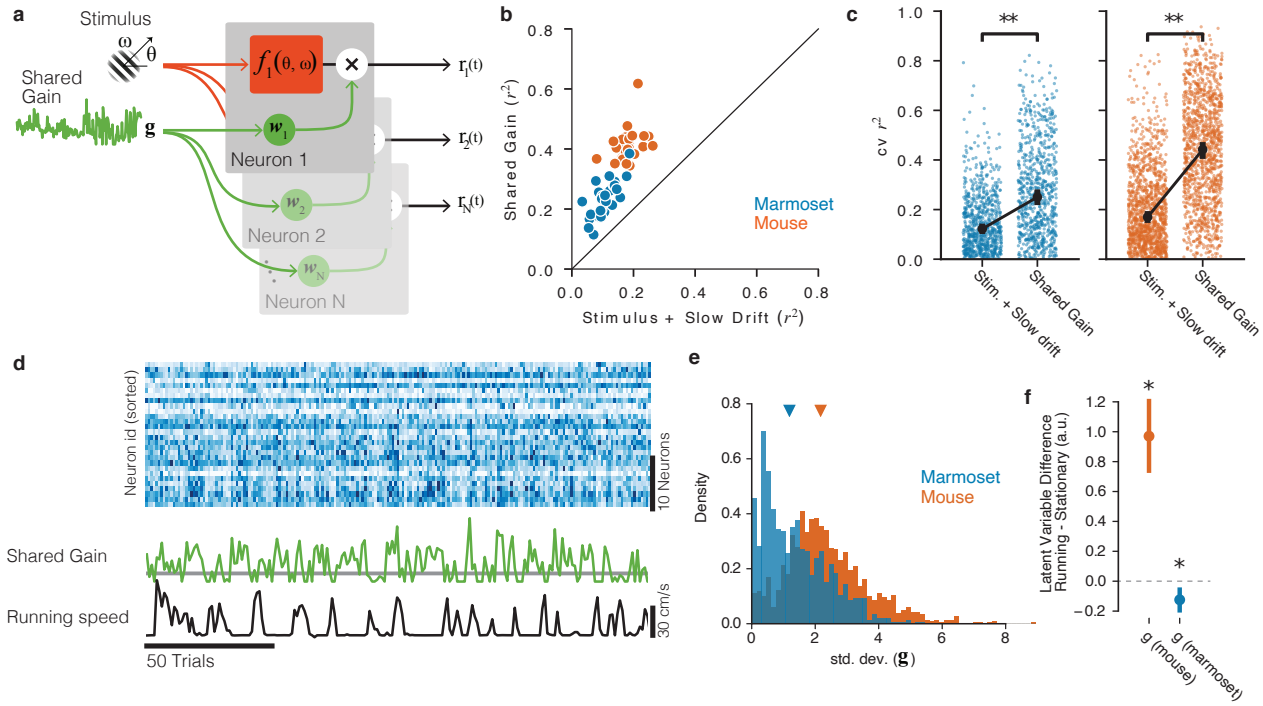

Figure 4. Shared gain model accounts for fluctuations in both mouse and marmoset V1, and explains species differences. a) Structure of shared modulator model. In addition to the effects of the stimulus (and slow drift in responsiveness, not rendered), the model allows for a shared gain/multiplicative term (green). Each simultaneously-recorded neuron is fitted with a weight to the latent gain term. b) The resulting model provides a better account of both mouse and marmoset V1 responses compared to a simple model that only fits stimulus and slow drift terms. Points show variance explained (r^2) on test data for each session under each of the two models, plotted against one another. c) Variance explained for individual units was significantly improved in both species (marmoset: gain model [median r^2=0.2504] significantly higher than stim+drift [median r^2 = 0.1220], p=1.52×10^{-82}, stat=27174, Wilcoxon signed rank test; mouse: gain model [median r^2=0.4420] significantly better than stim+drift [median r^2=0.1697], p=4.64×10^{-181}, stat=25966, Wilcoxon signed rank test). d) Example of relationship between neural responses (top raster, blue), the shared gain (green) and running speed (black trace). Visual inspection similar to that in Figure 2 can be performed. e) Gain modulations span a larger range in mice than in marmosets. Orange, gain term from each mouse session; blue, gain term from each marmoset session. Triangles indicate medians (mouse = 2.17 [2.11, 2.25], marmoset = 1.19 [1.07, 1.27]). f) Shared gain term is larger during running for mouse data, but is slightly smaller during running for marmoset data (difference is plotted on y-axis; mouse = 0.970 [0.761, 1.225], p=4.73×10^{-9}, stat 8.017, 1 sample t test; marmoset = -0.125 [-0.203, -0.059], p=0.002, stat=-3.360, 1 sample t test).

In summary, our results demonstrate that both mouse and marmoset V1 exhibit population-level gain modulations, likely reflecting modulatory inputs associated with behavioral state and arousal. Despite the commonality of mechanism, species (and likely, order-level) differences exist in how these modulations are linked to running, resulting in quantitative differences that are close to categorical. In mice, the large effects on V1 activity are likely to affect all subsequent stages of processing[13], but in marmosets, the small effects are less likely to have pronounced downstream effects. That said, running effects may occur at later stages of processing in primates, consistent with differences in where even canonical computations are located across species with different numbers of visual areas[15,39,40]. Likewise, finer-grained parcellation of what may be multiple effects correlated with running[3], and using cell-type-specific measures[1,4,8,9] will support more detailed comparisons across species and behavioral protocols. Of course, larger effects of behavioral state may still be found in primate V1: Other behaviors that better recruit active vision may reveal stronger modulations, and further consideration of the differences in feedback connections across the visual field representation[41] are now strongly motivated by our findings in peripheral V1. Exploring such possibilities can now enrich the cross-species connections we have drawn at the level of primate V1, as our identification of quantitative distinctions within a common mechanism lays a framework for continued cross-species generalizations that transcend simpler observations of empirical similarity or dissimilarity[42,43].

Methods

We performed electrophysiological recordings in V1 of two common marmosets (1 male, "marmoset G", and 1 female, "marmoset B", both aged 2 years). Both subjects had chronically implanted N-form arrays (Modular Bionics, Modular Bionics, Berkeley CA) inserted into left V1. Implantations were performed with standard surgical procedures for chronically-implanted arrays in primates. Additional recordings were also performed using Neuropixels 1.0 probes[44] acutely inserted into small craniotomies (procedure described below). All experimental protocols were approved by The University of Texas Institutional Animal Care and Use Committee and in accordance with National Institute of Health standards for care and use of laboratory animals.

Subjects stood quadrupedally on a 12" diameter wheel while head-fixed facing a 24" LCD (BenQ) monitor (resolution = 1920x1080 pixels, refresh rate = 120 Hz) corrected to have a linear gamma function, at a distance of 36 cm (pixels per degree = 26.03) in a dark room. Eye position was recorded via an Eyelink 1000 eye tracker (SR Research) sampling at 1 kHz. A syringe pump-operated reward line was used to deliver liquid reward to the subject. Timing events were generated using a Datapixx I/O box (VPixx) for precise temporal registration. All of these systems were integrated in and controlled by MarmoView. Stimuli were generated using MarmoView, custom code based on the PLDAPS[45] system using Psychophysics Toolbox[46] in MATLAB (Mathworks). For the electrophysiology data gathered from the N-Form arrays, neural responses were recorded using two Intan C3324 headstages attached to the array connectors which sent output to an Open Ephys acquisition board and GUI on a dedicated computer. In electrophysiology data gathered using Neuropixels probes, data was sent through Neuropixels headstages to a Neuropixels PXIe acquisition card within a PXIe chassis (National Instruments). The PXIe chassis sent outputs to a dedicated computer running Open Ephys with an Open Ephys acquisition board additionally attached to record timing events sent from the Datapixx I/O box. Spike sorting on data acquired using N-Form arrays was performed using in-house code to track and merge data from identified single units across multiple recording sessions[47]. Spike sorting for data acquired using Neuropixels probes was performed using Kilosort 2.5.

Chronic N-Form array recordings

Chronic array recordings were performed using 64-channel chronically-implanted 3D N-Form arrays consisting of 16 shanks arrayed in a 4x4 grid with shanks evenly spaced 0.4 mm apart (Modular Bionics, Berkeley, CA, USA). Iridium oxide electrodes are located at 1, 1.125, 1.25, and 1.5 mm (tip) along each shank, forming a 4x4x4 grid of electrodes. Arrays were chronically inserted into the left dorsal V1 of marmosets G and B at 1.5 and 4 degrees eccentric in the visual field, respectively (confirmed via post-hoc spatial RF mapping). Well-isolated single units were detectable on the arrays in excess of 6 months after the initial implantation procedure.

Acute Neuropixels recordings

Acute Neuropixels recordings were performed using standard Neuropixels 1.0 electrodes (IMEC, Leuven, Belgium). Each probe consists of 384 recording channels that can individually be configured to record signals from 960 selectable sites along a 10-mm long, 70 x 24 µm cross-section straight shank. Probes were lowered into right dorsal V1 of marmoset G via one of 3 burr holes spaced irregularly along the AP axis 4-5 mm from the midline for a single session of experiments. Natural images were played to provide visual stimulus as well as occupy the subject and keep them awake during insertion and probe settling. The temporary seal on the burr hole was removed, the intact dura nicked with a thin needle and the burr hole filled with saline. The probe was then lowered through the dural slit at 500 µm/minute, allowing 5 minutes for settling every 1000 µm of total insertion. The whole-probe LFP visualization was monitored during insertion for the characteristic banding of increased LFP amplitude that characterizes cortical tissue. The probe was inserted until this banding was visible on the electrodes nearest the tip of the probe, indicating that the probe tip itself had passed through the dorsal cortex and was within the white matter. The probe was then advanced until a second band became visible on the electrodes nearest the tip, indicating the tip of the probe had exited through the cortex of the calcarine sulcus. The probe was then advanced slightly until the entirety of the second LFP band was visible to ensure that electrodes covered the full depth of the calcarine cortex and the tip of the probe was located confidently within the CSF of the sulcus. The probe was then allowed to settle for 10 minutes. Active electrode sites on the probe were configured to subtend both dorsal and calcarine cortex simultaneously. Post-hoc receptive field recreation confirmed that visually-driven, tuned, V1 neurons were recorded at both foveal and peripheral eccentricities.

Mouse dataset from Allen Institute

Mouse data were downloaded from the publicly-available Visual Coding database at https://portal.brain-map.org/explore/circuits/visual-coding-neuropixels. We used the same analysis code to analyze these data and the marmoset data we collected.

Experimental workflow

Marmoset recording sessions began with eye tracking calibration. Once calibration was completed, the wheel was unlocked and the subject was allowed to locomote freely, head-fixed, while free-viewing stimuli. Trials for all stimuli were 20 sec long with a

500 ms ITI and a 20 sec long natural image interleaved every fifth trial to keep the subject engaged. Stimuli were shown in blocks of 10 minutes and a typical recording session consisted of 50 trials of calibration followed by 1 or 2 blocks of a drifting grating stimulus and 1 block each of the two mapping stimuli. To elicit sufficiently reliable and frequent running behavior, subjects were rewarded at set locomotion distance intervals unrelated to the stimulus or gaze behavior (typical rewards were 50-70 μL and distance required to achieve a reward usually varied between 20-75 cm; reward amounts and intervals were adjusted daily to maximally motivate the subject.)

Eye tracking calibration

While the wheel was locked, subjects were allowed to free-view a sequence of patterns of marmoset faces. Marmosets naturally direct their gaze towards the faces of other marmosets when allowed to free-view with little-to-no training, allowing for the experimenter to adjust the calibration offset and gain manually between pattern presentations. Faces were 1.5 degrees in diameter and were presented for 3 sec with a 2 sec ISI between patterns. A portion of presented patterns were asymmetrical across both the X and Y axes of the screen to allow for disambiguation in the case of axis sign flips in the calibration. 50 trials were presented before each recording session to verify and refine the calibration. Calibration drift between sessions was minimal, requiring minor (<1 deg) adjustments over the course of 1-2 months of recordings.

Drifting grating stimuli

The primary stimulus consisted of full-field drifting gratings. Gratings were optimized to drive marmoset V1 with 3 separate spatial frequencies (1, 2, and 4 cycles per degree), two drift speeds (1 or 2 degrees per second) and 12 orientations (evenly-spaced 30 degree intervals). Each trial consisted of multiple grating presentations, each with a randomized spatial frequency, drift speed, and orientation. Gratings were displayed for 833 ms followed by a 249-415 ms randomly jittered inter-stimulus interval. After each 20 second trial there was a longer 500 ms inter-trial interval. Every fifth trial was replaced with a natural image to keep subjects engaged and allow for visual assessment of calibration stability on the experimenter's display.

Mapping of receptive fields

A spatiotemporal receptive field mapping stimulus, consisting of sparse dot noise, was shown during each recording session. One hundred 1 degree white and black dots were presented at 50% contrast at random points on the screen. Dots had a lifetime of 2 frames (16.666 ms). Marmosets freely viewed the stimulus and we corrected for eye position offline to estimate the spatial receptive fields using forward correlation[20].

Necessary differences between mouse and marmoset experiments

Although we sought to perform experiments in marmosets that were as similar as possible to mouse experiments, some differences in their visual systems and behavior made for differences. Because the spatial frequency tunings of marmoset and mouse V1 neurons are starkly different, we used stimuli with considerably higher spatial frequencies than in the mouse experiments. Relatedly, marmoset V1 receptive fields are much smaller than in mouse. Because we used full-field stimuli (to match mouse experiments), responses in marmoset V1 were likely affected by substantial amounts of surround suppression, which would reduce overall responses. We also learned that, although the marmosets were comfortable perched on the wheel treadmill, they did not naturally run enough for our experimental purposes. We therefore incorporated a reward scheme to motivate the subjects to run more frequently. Finally, the mouse dataset we analyzed comprised a large number of mice with a small number of sessions per mouse; as is required of work with nonhuman primates, we were limited to a smaller number of subjects (N=2), and ran many experimental sessions with each animal.

Session and cell inclusion criteria

For the analyses shown in Figure 2, sessions were included if they contained more than 250 trials and a proportion of trials running was not less than 10% or greater than 90%. For the mouse dataset, this yielded 25/32 sessions. For the marmoset dataset, this yielded 27/34 sessions. For the unit-wise analyses in Figure 3, super-sessioned units were included for analysis if they had more than 300 trials of data and a mean firing rate of >1 spike / second. This yielded 1168/2015 units in mouse and 786/1837 units in marmoset.

For the analyses shown in Figure 4, sessions were included using the same trial and running criterion as in Figure 2. Only units that were well fit by the stimulus + slow drift model (i.e., cross validated better than the null, see 'shared modulator model') were included and sessions were excluded if fewer than 10 units met this criterion. This resulted in 31/32 sessions for mouse and 28/34 sessions for marmoset.

Analysis of tuning

We counted spikes between the 50ms after grating onset and 50ms after grating offset and divided by the interval to generate a trial spike rate. To calculate orientation tuning curves, we computed the mean firing rate each orientation and spatial frequency.

Because we were limited by the animal's behavior to determine the number of trials in each condition (i.e., running or not), we computed orientation tuning as a weighted average across spatial frequencies with with weights set by the spatial frequency tuning. We used these resulting curves for the all analyses of tuning. We confirmed that the results did not change qualitatively if we either used only the best spatial frequency or marginalized across spatial frequency.

Orientation selectivity index was calculated using the following equation

$$\mathrm{OSI} = \frac{\sqrt{[r^T \sin(2\theta)]^2 + [r^T \cos(2\theta)]^2}}{\sum(r)}$$

where θ is the orientation and r is the baseline-subtracted vector of rates across orientations.

Shared modulator model

To capture shared modulator signals in an unsupervised manner, we fit our neural populations with a latent variable model[48]. The goal of our latent variable model is to summarize population activity with low-dimensional shared signal that operates as a gain on the stimulus processing (e.g.[24,26]). The general form of the model is that the response of an individual neuron, $\mathbf{r}_i$ on trial t is decomposed into a stimulus response, gain modulator, and additive offsets:

$$\mathbf{r}_i(t) = f_i[s(t)] * g_i(t) + b_i \tag{1}$$

where $f_i[s(t)]$ is the tuning curve, $g_i(t)$ is a neuron-specific gain on the stimulus response, $h_i(t)$ is an additive noise term for the trial and b_i is the baseline firing rate. To scale this to a shared population model, we enforced the gain, g, to be rank 1, such that it can be decomposed into a trial-wise vector of gains and a neuron-wise vector of loadings that map the trial latent into modulatory signal for each neuron. Similar models have been employed to describe the population response in V1 in several species[23–26].

To capture the stimulus tuning curves, we represented the stimulus on each trial an $m-$dimensional "one-hot" vector, where m is the number of possible conditions (Orientation $\times$ Spatial Frequency) and on each trial all elements are zero, except for the condition shown. Thus, $f[s(t)]$ is a linear projection of the stimulus on the tuning curves, $\mathrm{A}s(t)$, where W is an $n \times m$ matrix of tuning weights. We decomposed the gain for each neuron on each trial into a rank 1 matrix that was rectified and offset by one, $g(t) = \mathrm{ReLU}[1 + \mathrm{z_g(t)w_g}]$, where w_g is an $n-$dimensional vector of loadings that map the 1-dimensional trial latent to a population-level signal, $z_g(t)\mathbf{w}_g$. This signal is offset by 1 and rectified such that it is always positive and a loading weight of zero equals a gain of 1.0.

Thus, the full model describes the population response as

$$\mathbf{r}(t) = \mathbf{A}\mathrm{s}(t)\mathrm{F}[1 + g(t)\mathbf{w}_g] + \mathbf{b} \tag{2}$$

Thus, the parameters of the model are the stimulus tuning parameters $\mathbf{A}$, the shared gain, g, the gain loadings, $\mathbf{w}_g$, and the offsets, $\mathbf{b}$. To capture any unit-specific slow drifts in firing rate, we further parameterized $\mathbf{b}$ as a linear combination of 5 b0-splines evenly spaced across the experiment[49]. Thus, the baseline firing rate for each neuron, i, was a linear combination of 5 "tent" basis functions spaced evenly across the experiment, $\mathbf{b}_i = \sum_j b_j \phi_j(t)$.

We first fit a baseline model with only stimulus and baseline parameters

$$\mathbf{r}(t) = \mathbf{A}\mathrm{s}(t) + \mathbf{b} \tag{3}$$

Following Whiteway and Butts (2017), we initialized $\mathbf{A}$ and $\mathbf{b}$ the model using fits from a model without latent variables and initialized the latent variables using an Autoencoder[50,51]. We then fit the gain, loadings, and stimulus parameters using iterative optimization with L-BFGS, by minimizing the mean squared error (MSE) between the observed spikes and the model rates. The model parameters were regularized with a modest amount of L2-penalty and the amount was set using cross-validation on the training set. The latent variables were penalized with a small squared derivative penalty to impose some smoothness across trials. This was set to be small and the same value across all sessions. We reverted the model to the autoencoder initialization if the MSE on a validation set did not improve during fitting.

We cross-validated the model using a speckled holdout pattern[52] whereby some fraction of neurons were withheld on each trial with probability p=0.25. We further divided the withheld data into a validation set and a test set by randomly assigning units to either group on each trial with probability 0.5. The validation loss was used to stop the optimization during the iterative fitting and the test set was used to evaluate the models.

Acknowledgements

We thank Allison Laudano for animal and colony management and care, Christopher Badillo for apparatus design and fabrication, and Nika Hazen for assistance with animal work. Cris Niell, Cory Miller, and Jude Mitchell all provided valuable feedback on drafts of this paper. This work was supported by U01-UF1NS116377 to A.H. NSF NSC-FO 2123605 (to D.B.) and NIH K99EY032179-02 (to J.Y.).

Author contributions statement

A.H. and J.L. conceived the experiment(s), J.Y. and D.B. conceived the analyses, J.L., D.R., and T.N. conducted the experiment(s), J.Y., D.B., J.L, and D.R analysed the results. All authors wrote and reviewed the manuscript.

References

1. Niell, C. M. & Stryker, M. P. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron* **65**, 472–479 (2010).
2. Keller, G. B., Bonhoeffer, T. & Hübener, M. Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. *Neuron* **74**, 809–815 (2012).
3. Vinck, M., Batista-Brito, R., Knoblich, U. & Cardin, J. A. Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. *Neuron* **86**, 740–754 (2015).
4. Bennett, C., Arroyo, S. & Hestrin, S. Subthreshold Mechanisms Underlying State-Dependent Modulation of Visual Responses. *Neuron* **80**, 350–357 (2013).
5. Saleem, A. B., Ayaz, A., Jeffery, K. J., Harris, K. D. & Carandini, M. Integration of visual motion and locomotion in mouse visual cortex. *Nat. Neurosci.* **16**, 1864–1869 (2013).
6. Erisken, S. *et al.* Effects of Locomotion Extend throughout the Mouse Early Visual System. *Curr. Biol.* **24**, 2899–2907 (2014).
7. Reimer, J. *et al.* Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. *Neuron* **84**, 355–362 (2014).
8. Pakan, J. M. *et al.* Behavioral-state modulation of inhibition is context-dependent and cell type specific in mouse visual cortex. *eLife* **5**, e14985 (2016).
9. Polack, P.-O., Friedman, J. & Golshani, P. Cellular mechanisms of brain state–dependent gain modulation in visual cortex. *Nat. Neurosci.* **16**, 1331–1339 (2013).
10. Fu, Y. *et al.* A Cortical Circuit for Gain Control by Behavioral State. *Cell* **156**, 1139–1152 (2014).
11. Mineault, P. J., Tring, E., Trachtenberg, J. T. & Ringach, D. L. Enhanced Spatial Resolution During Locomotion and Heightened Attention in Mouse Primary Visual Cortex. *J. Neurosci.* **36**, 6382–6392 (2016).
12. Ayaz, A., Saleem, A. B., Schölvinck, M. L. & Carandini, M. Locomotion Controls Spatial Integration in Mouse Visual Cortex. *Curr. Biol.* **23**, 890–894 (2013).
13. Christensen, A. J. & Pillow, J. W. Reduced neural activity but improved coding in rodent higher-order visual cortex during locomotion. *Nat. Commun.* **13**, 1676 (2022).
14. Rosa, M. G. P. & Krubitzer, L. A. The evolution of visual cortex: where is V2? *Trends Neurosci.* **22**, 242–248 (1999).
15. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex (New York, N.Y.: 1991)* **1**, 1–47 (1991).
16. Niell, C. M. & Stryker, M. P. Highly Selective Receptive Fields in Mouse Visual Cortex. *J. Neurosci.* **28**, 7520–7536 (2008).
17. Hoy, J. L., Yavorska, I., Wehr, M. & Niell, C. M. Vision Drives Accurate Approach Behavior during Prey Capture in Laboratory Mice. *Curr. Biol.* **26**, 3046–3052 (2016).
18. Land, M. F. & Fernald, R. D. The evolution of eyes. *Annu. Rev. Neurosci.* **15**, 1–29 (1992).
19. Visual Coding - Neuropixels - brain-map.org.
20. Yates, J. L. *et al.* Beyond Fixation: detailed characterization of neural selectivity in free-viewing primates. Tech. Rep., bioRxiv (2021).
21. Yu, H.-H. & Rosa, M. G. P. Uniformity and diversity of response properties of neurons in the primary visual cortex: Selectivity for orientation, direction of motion, and stimulus size from center to far periphery. *Vis. Neurosci.* **31**, 85–98 (2014).
22. Sengpiel, F., Troilo, D., Kind, P. C., Graham, B. & Blakemore, C. Functional architecture of area 17 in normal and monocularly deprived marmosets (Callithrix jacchus). *Vis. Neurosci.* **13**, 145–160 (1996).
23. Whiteway, M., Butts, D., Socha, K. & Bonin, V. Characterizing the nonlinear structure of shared variability in cortical neuron populations using latent variable models. *Neurons, Behav. Data analysis, Theory* **2**, 1–22 (2019).

24. Lin, I.-C., Okun, M., Carandini, M. & Harris, K. D. The Nature of Shared Cortical Variability. *Neuron* **87**, 644–656 (2015).

25. Arandia-Romero, I., Tanabe, S., Drugowitsch, J., Kohn, A. & Moreno-Bote, R. Multiplicative and Additive Modulation of Neuronal Tuning with Population Activity Affects Encoded Information. *Neuron* **89**, 1305–1316 (2016).

26. Goris, R. L. T., Movshon, J. A. & Simoncelli, E. P. Partitioning neuronal variability. *Nat. Neurosci.* **17**, 858–865 (2014).

27. Ferguson, K. A. & Cardin, J. A. Mechanisms underlying gain modulation in the cortex. *Nat. Rev. Neurosci.* **21**, 80–92 (2020).

28. Disney, A. A. & Robert, J. S. Translational implications of the anatomical nonequivalence of functionally equivalent cholinergic circuit motifs. *Proc. Natl. Acad. Sci.* **116**, 26181–26186 (2019).

29. Coppola, J. J. & Disney, A. A. Is There a Canonical Cortical Circuit for the Cholinergic System? Anatomical Differences Across Common Model Systems. *Front. Neural Circuits* **12** (2018).

30. Disney, A. A., Aoki, C. & Hawken, M. J. Gain Modulation by Nicotine in Macaque V1. *Neuron* **56**, 701–713 (2007).

31. Lien, A. D. & Scanziani, M. Tuned thalamic excitation is amplified by visual cortical circuits. *Nat. Neurosci.* **16**, 1315–1323 (2013).

32. Pfeffer, C. K., Xue, M., He, M., Huang, Z. J. & Scanziani, M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* **16**, 1068–1076 (2013).

33. Chen, X., Cao, L. & Haendel, B. F. Differential effects of walking across visual cortical processing stages. *Cortex* **149**, 16–28 (2022).

34. Cao, L., Chen, X. & Haendel, B. F. Overground Walking Decreases Alpha Activity and Entrains Eye Movements in Humans. *Front. Hum. Neurosci.* **14** (2020).

35. Benjamin, A. V., Wailes-Newson, K., Ma-Wyatt, A., Baker, D. H. & Wade, A. R. The Effect of Locomotion on Early Visual Contrast Processing in Humans. *J. Neurosci.* **38**, 3050–3059 (2018).

36. Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* **63**, 879–888 (2009).

37. Cohen, M. R. & Maunsell, J. H. R. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* **12**, 1594–1600 (2009).

38. Rabinowitz, N. C., Goris, R. L., Cohen, M. & Simoncelli, E. P. Attention stabilizes the shared gain of V4 populations. *eLife* **4**, e08998 (2015).

39. 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).

40. Scholl, B., Tan, A. Y. Y., Corey, J. & Priebe, N. J. Emergence of Orientation Selectivity in the Mammalian Visual Pathway. *J. Neurosci.* **33**, 10616–10624 (2013).

41. Wang, M. *et al.* Retinotopic organization of feedback projections in primate early visual cortex: implications for active vision. preprint, Neuroscience (2022).

42. Niell, C. M. & Scanziani, M. How Cortical Circuits Implement Cortical Computations: Mouse Visual Cortex as a Model. *Annu. Rev. Neurosci.* **44**, 517–546 (2021).

43. Priebe, N. J. & McGee, A. W. Mouse vision as a gateway for understanding how experience shapes neural circuits. *Front. Neural Circuits* **8** (2014).

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

45. Eastman, K. M. & Huk, A. C. PLDAPS: A Hardware Architecture and Software Toolbox for Neurophysiology Requiring Complex Visual Stimuli and Online Behavioral Control. *Front. Neuroinformatics* **6** (2012).

46. Brainard, D. H. The Psychophysics Toolbox. *Spatial Vis.* **10**, 433–436 (1997).

47. Muthmann, J.-O., Levi, A. J., Carney, H. C. & Huk, A. C. A hardware/software system for electrophysiology "supersessions" in marmosets. Tech. Rep., bioRxiv (2021).

48. Whiteway, M. R. & Butts, D. A. The quest for interpretable models of neural population activity. *Curr. Opin. Neurobiol.* **58**, 86–93 (2019).

49. Quinn, K. R., Seillier, L., Butts, D. A. & Nienborg, H. Decision-related feedback in visual cortex lacks spatial selectivity. *Nat. Commun.* **12**, 4473 (2021).

50. Bengio, Y., Courville, A. & Vincent, P. Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis Mach. Intell.* **35**, 1798–1828 (2013).

51. Whiteway, M. R. & Butts, D. A. Revealing unobserved factors underlying cortical activity with a rectified latent variable model applied to neural population recordings. *J. Neurophysiol.* **117**, 919–936 (2017).

52. Williams, A. H. *et al.* Unsupervised Discovery of Demixed, Low-Dimensional Neural Dynamics across Multiple Timescales through Tensor Component Analysis. *Neuron* **98**, 1099–1115.e8 (2018).