Guided by Noise: Correlated Variability Channels Task-Relevant Information in Sensory Neurons

Ramanujan Srinath[1], Yunlong Xu[2], Douglas A. Ruff[1], Amy M. Ni[1], Brent Doiron[1,2,3], Marlene R. Cohen[1]

[1]Department of Neurobiology and Neuroscience Institute, The University of Chicago, Chicago, IL 60637, USA
[2]Grossman Center for Quantitative Biology and Human Behavior, University of Chicago, Chicago, IL, USA
[3]Department of Statistics, University of Chicago, Chicago, IL, USA

Abstract

Shared trial-to-trial variability across sensory neurons is reliably reduced when perceptual performance improves, yet this variability is low-dimensional, so it could be ignored by an optimal readout mechanism. Why then is it so consistently related to behavior? We propose that ***shared variability reflects the information communicated to downstream areas, rather than noise to be avoided***. Using a circuit model, we show that when sensory signals align with shared variability, behaviorally relevant information is amplified without compromising coding fidelity. Analyses of neural population recordings from multiple brain areas and tasks reveal that the dominant axis of *shared variability consistently aligns with task-relevant stimulus features and action plans*. Finally, the behavioral impact of microstimulation can be explained by the extent to which it changes projections onto the shared variability axis. These findings suggest that shared variability may illuminate, rather than obscure, the neural dimensions that guide behavior.

Introduction

Correlated neural variability is both tightly linked to behavior and, under some encoding frameworks, irrelevant to information coding. Many studies have shown that good performance on perceptual tasks, associated with factors such as attention, learning, arousal/motivation, or the contrast of a visual stimulus, is consistently accompanied by low shared trial-to-trial response fluctuation in sensory neurons[1–15]. Shared variability is often quantified as the mean spike count correlation (also called noise correlations or r_{SC}) between the trial-to-trial fluctuations in the responses of pairs of neurons to repeated presentations of the same stimulus[1]. Shared variability has become central to experimental and theoretical investigations of information coding, in part because it is modulated by virtually every process known to enhance perception[1,9,16], linked to behavior on a trial-by-trial basis[17–20], and sensitive to pharmacological and circuit-level manipulation[5,21,22].

These observations initially fueled a widely held hypothesis: that shared variability acts as noise, degrading the fidelity of population codes and impairing the ability of downstream areas to make accurate perceptual decisions[23]. From this perspective, cognitive processes such as attention or arousal improve perception by reducing shared noise[24,25]. However, this interpretation is increasingly difficult to reconcile with theoretical and empirical evidence. Because correlated variability tends to lie within a low-dimensional subspace of neural activity (defined as the space in which the response of each neuron represents one dimension[16,23–27]), an optimal decoder can, in principle, ignore it entirely[16,23,26–29]. Consistent with this, we recently demonstrated that attention dramatically reduces correlated variability in the visual cortex without substantially altering the amount of sensory information that can be linearly decoded from the population[30]. Moreover, the same axis that captures the most shared variability also explains nearly all of the task-related choice signals in visual area V4[31]. These observations present a paradox: if correlated variability does not fundamentally constrain the information available in a sensory population, why is it so reliably related to behavior?

Here, we propose a resolution to this contradiction. Rather than interpreting shared variability as an obstacle to neural coding, we suggest that it reflects the dimensions of neural population activity that guide behavior. Specifically, we hypothesize that behaviorally relevant information is preferentially aligned with the axis of shared variability (correlated variability axis), defined as the axis in the neural population along which responses fluctuate most across trials. We first demonstrate support for our hypothesis using previously collected V4 population recordings during a change detection task. In this dataset, detection performance was higher when neuronal responses were better aligned with the correlated variability axis. Motivated by this initial support, we next sought a theoretical rationale for this alignment by studying stimulus estimation in a recurrently coupled network of excitatory and inhibitory neuron models subject to external noise and a separate input signal. The recurrent coupling defines a one-dimensional correlated variability axis, and indeed,

when signals are aligned with that axis, that same recurrent circuitry also amplifies the response to the signal, so that overall information available to a (linear) decoder is maximized. We tested this general principle across multiple datasets by analyzing how correlated variability relates to representations of relevant and irrelevant stimulus features, motor plans, and decisions. In a curvature estimation task, the axis of correlated variability was aligned with the representation of curvature across heterogeneous shapes, irrespective of irrelevant features. During saccade planning, correlated variability more closely tracked motor plans than sensory inputs. In a task where animals alternated discriminating curvature and color, the alignment of curvature information with the correlated variability axis increased when that feature was relevant to choices. Finally, we found causal support for this framework: the behavioral impact of microstimulation was predicted by how closely evoked population responses aligned with the axis of correlated variability.

Together, these findings challenge the idea that neural representations avoid subspaces of neural population space corrupted by correlated variability and instead suggest that behaviorally relevant signals are communicated along these subspaces. By reframing the role of shared variability, our findings provide a new perspective on decades of experimental results, suggesting that correlated variability may serve as a window into the neural computations that drive perception and decision-making.

Results

While improvements in perceptual performance often coincide with changes in correlated trial-to-trial variability in sensory neurons, theoretical work suggests that such variability may not limit the information that is available to optimal decoders[16,23]. This is because correlated variability, especially in sensory areas, typically occupies a low-dimensional subspace, fluctuating primarily along a single “correlated variability axis”, and decoders can readily organize to be orthogonal to this axis, making them blind to shared fluctuations. We investigated our central hypothesis by analyzing the relationship between task-relevant information coding in the visual cortex, correlated variability, and behavior. In each data set, we defined the axis of correlated variability as the first principal component (PC1) of neural population activity[31,32] during a baseline period where rhesus monkeys (Macaca mulatta) fixated a gray screen at the start of each trial, capturing the dimension that explains the most shared variability in spontaneous activity.

Performance in an orientation change detection task is related to the alignment between the orientation representation and the axis of correlated variability in area V4

We began by testing our hypothesis in a well-established orientation change detection task[2,9,12,20,30,31,33] (Figure 1A). Different aspects of these data have been published previously[21,31]. In this task, monkeys fixated a central dot while two Gabor stimuli of the same orientation were flashed repeatedly. At a random and unsignalled time, one stimulus changed orientation, and the monkey was rewarded for making a saccade to its location. In this task variation, the starting orientations of the gratings varied between 0° and 180° across trials, while the magnitude of the orientation change was fixed. Our analysis focused on trials in which the change occurred at a cued location that overlapped with the joint receptive fields of the recorded V4 neurons.

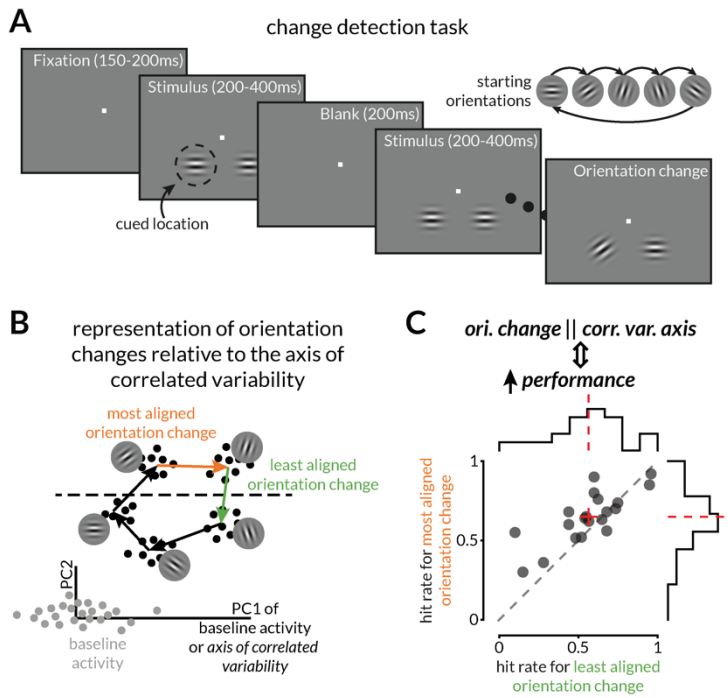

Figure 1: Behavioral performance in an orientation change detection task is related to the alignment between V4 orientation representations and the axis of correlated variability.

A. Schematic of the orientation change detection task. Two Gabor stimuli with the same initial orientation are flashed repeatedly, and one changes orientation at a random time picked from an exponential distribution. Monkeys are rewarded for making a saccade to the location of the orientation change within 500 ms[9,31]. On each trial, the initial orientation was randomly selected between 0° and 180°, while the magnitude of the change was constant across trials. The stimulus most likely to change was cued in blocks (80% valid). We analyzed trials in which the change occurred at the cued location when that location was within the joint receptive fields of the recorded V4 neurons.
B. We defined the *axis of correlated variability* as the first principal component of V4 activity during the period where the monkey fixated a gray screen before the first stimulus presentation (gray dots represent these baseline responses). Black dots represent evoked responses to stimuli with different orientations. Our hypothesis predicts that performance is better when the orientation change aligns with the axis of correlated variability (most aligned change; orange arrow) than when it does not (least aligned orientation change; green arrow).
C. Validation of the prediction in B. Each point represents the comparison of the average hit rates on the most and least aligned orientation change conditions for one session (n=20; approx. 426 trials per session; approx. 18 trials per orientation change at the attended location). Across sessions, monkeys performed better on the best aligned than on the least aligned orientation change ($p<0.05$; Wilcoxon signed-rank test). Marginal histograms show the distribution of hit rates across sessions. The red plus sign and red dashed lines denote man values.

According to our hypothesis, behavioral performance should be better for orientation changes that are most aligned with the axis of correlated variability (Figure 1B, orange arrow) compared to orientation changes that are least aligned (Figure 1B, green arrow). We quantified alignment as the performance of a leave-one-out cross-validated classifier trained on the neural responses projected onto the axis of correlated variability to classify the starting and change orientations. Consistent with this prediction, monkeys were significantly better at detecting the orientation change most aligned with the axis of correlated variability compared to the least aligned orientation change (Figure 1C). This unexpected result in a well-studied task prompted us to explore our hypothesis further using a theoretical model and four additional datasets, which encompassed a wide range of stimuli and behavioral demands.

A circuit model demonstrates that stimulus information can be read out optimally when aligned with the axis of correlated variability

To explore the neuronal circuitry required to align stimulus representations with the correlated variability axis, we constructed a recurrently coupled network model composed of excitatory and inhibitory units with defined feedforward (W_F), recurrent (W_R), and linear read-out (W_O) synaptic weights (Figure 2A). In our model, each neuron receives an external source of private (independent) trial-to-trial fluctuations, so that population co-variability only emerges through recurrent interactions between neurons (as defined by W_R). We adopt this formulation because, across all datasets, the axis of correlated variability measured during repeated presentations of the same stimulus is nearly parallel to the axis obtained during baseline gray-screen periods (Figure S1), indicating that recurrent connectivity is the dominant source of these correlations.

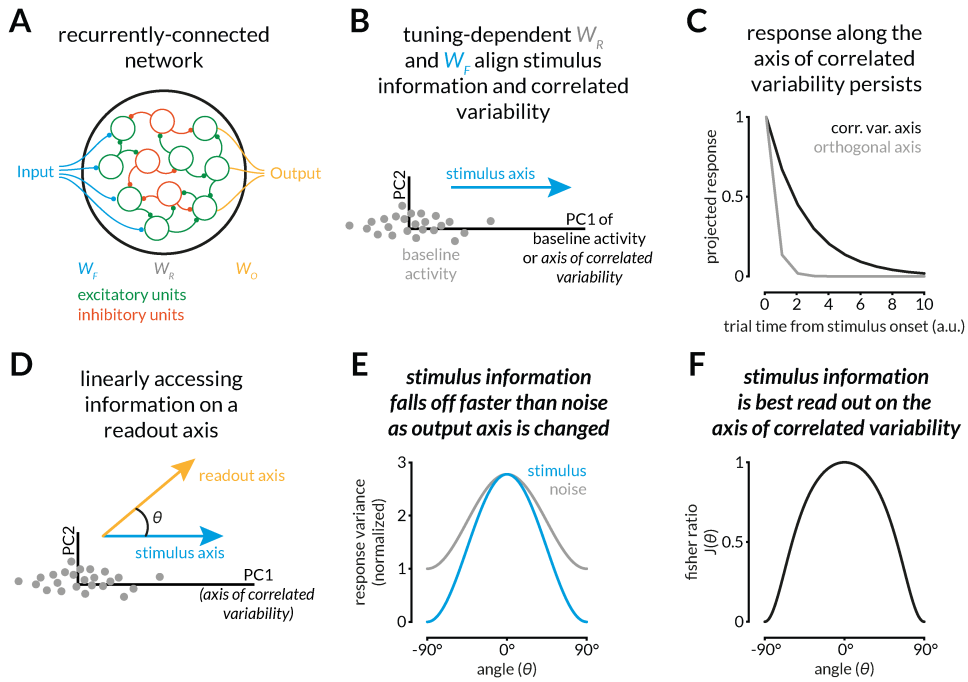

Figure 2. Alignment of stimulus information with the axis of correlated variability makes it optimal for readout.

A. Recurrently coupled network composed of excitatory and inhibitory neuron models with rank-one feed-forward weights W_F (blue), recurrent weights W_R (grey), and linear readout weights W_O (orange). A stimulus enters through W_F. Independent private noise is injected into each neuron and is shaped only by recurrent weights (W_R). The network connectivity follows Dale's law.
B. In our network, since the recurrent connectivity (W_R) is aligned with the feedforward input (W_F), the stimulus axis (blue arrow) aligns with the first principal component of baseline activity (PC1, grey dots)
C. Once this alignment is established, fluctuations along PC1 decay much more slowly than along any orthogonal mode. Power on the slow mode (black curve) persists, whereas power on, for example, PC2 (light grey) vanishes rapidly. Assigning the task-relevant input to the slowest decaying dimension enables the circuit to integrate information over time, providing a normative rationale for the recurrent tuning.
D. We examine a linear readout whose axis (orange) forms an angle θ with the stimulus axis (blue).
E. Normalized information about the stimulus (blue) and noise (gray) as a function of θ. As the read-out is rotated away from the axis of correlated variability, stimulus information decreases more steeply than noise.
F. Fisher discriminability $\mathrm{J}(\theta) = \frac{\mathrm{signal}(\theta)}{\mathrm{noise}(\theta)}$ peaks at $\theta = 0°$, demonstrating that, when the stimulus and noise axes are aligned, the optimal linear decoder is along the noise axis.

A wealth of data demonstrates that pairwise signal correlations are related to the magnitude of trial-to-trial pairwise noise correlations; neurons with similar tuning have stronger noise correlations[1,34,35]. A likely contributor to this relationship is that recurrent coupling (W_R) is stimulus feature dependent (W_F), as reported in mouse V1[36,37]. Therefore, we assume that the recurrent connectivity is tuned and build W_R directly from W_F (see Methods). Additionally, across our datasets, the axes of correlated variability computed from the baseline period and from repeated presentations of the same stimulus are well aligned (Figure S1). This independence of correlated variability from the stimulus suggests that it is not strongly shaped by feedforward inputs. We therefore model correlated variability as noise that is shaped only by W_R, but not by W_F.

The relationship between W_R and W_F creates a biologically realistic alignment between stimulus-evoked activity and the axis of correlated variability (Figure 2B; Supplementary Information Appendix A). Appendix A shows that if the recurrent weights are unrelated to stimulus feature tuning, i.e., arbitrary and rank-one, this alignment is not guaranteed. Instead, this alignment comes from the relationship between W_R and W_F in our model. We also observe that after the onset of the stimulus on a given trial, response fluctuations along the axis of correlated variability persist for a longer amount of time within the trial (Figure 2C, black curve) compared to fluctuations along orthogonal axes (Figure 2C, light gray curve). For an unconstrained one-dimensional decoder, a balanced change in signal and noise leaves the Fisher information unchanged, independent of the angle between the stimulus response axis and the correlated-variability axis (Supplementary Information Appendix C.2). By contrast, if the decoder is locked to a particular network mode, the alignment between stimulus-evoked activity and the axis of correlated variability increases information (Supplementary Information Appendix C.3).

We next explored how the alignment between stimulus and readout axes impacts decoding. To measure this in our model, we assumed that downstream areas can access only a single dimension of population activity (the readout axis) and that the corresponding readout weights (W_O) remain fixed throughout the task. This is consistent with evidence that cortical circuits exploit privileged axes embedded in hard-wired projection[38]. We changed the angle θ between the linear readout axis (Figure 2D, orange arrow) and the stimulus representation (blue arrow), and measured how much signal and noise information is aligned with the readout axis. As this angle increases, the normalized signal variance falls more rapidly than the noise variance (Figure 2E). Consequently, as predicted in the proofs in Supplementary Information Appendix B, the Fisher discriminability, a measure of decoding performance, peaks when the readout and noise axes are aligned (Figure 2F). These results demonstrate that when stimulus information is aligned with the axis of correlated variability (a biologically relevant scenario, as correlated variability is related to signal in essentially every study that measures both[1,10,39–45]), the optimal readout strategy is along the axis of correlated variability. The optimality of the noise axis for decoding is proven formally in Supplementary Information Appendix B.

Our model predicts that aligning behavior with the axis of correlated variability is optimal when the relevant signal is aligned with noise. To achieve optimal coding, the core assumption of our model is that recurrent synaptic interactions (that determine the coverability axis) align with the feedforward synaptic weights (that carry signal information). This is a general assumption, consistent with the pairwise signal and noise correlations themselves being correlated, as widely reported in the literature[2,43,46].

Alignment of curvature representation with the axis of correlated variability in V4 predicts behavioral performance in a curvature estimation task

We next tested the prediction that, in a situation where monkeys had to generalize across stimuli that varied in many irrelevant features, the axis of correlated variability would align with the axis that best facilitates that generalization. We trained monkeys to perform a continuous curvature estimation task (Figure 3A), in which they reported the curvature of any random 3D shape while ignoring other irrelevant features, such as color, orientation, and thickness profile[48,47]. After the stimulus was presented, the monkeys made a saccade to a location on a target arc corresponding to their curvature estimate (0 for straight, 1 for maximally curved). They were rewarded in inverse proportion to their estimation error.

Both monkeys (Figure 3B, green and blue lines) exhibited consistent curvature estimation, with their performance comparable to that of humans performing an online version of the same task (Figure 3B, pink line). The stimuli were presented within the joint receptive fields of recorded V4 neurons, which are selective for curvature (Figure 3D-E). The population representation of curvature was dependent on irrelevant shape properties (Figure 3F-G)[47], meaning that the best axis for judging the curvature of any one stimulus was different than the best for generalizing across stimuli.

Our central hypothesis predicts that behavioral performance should be better for shapes whose curvature representation is more aligned with the axis of correlated variability (Figure 3C). To test this, we compared the behavioral performance across pairs of shapes tested within the same experimental session (Figure 3H). Consistent with our prediction, performance was consistently better for the shape whose curvature representation was better aligned with the axis of correlated variability.

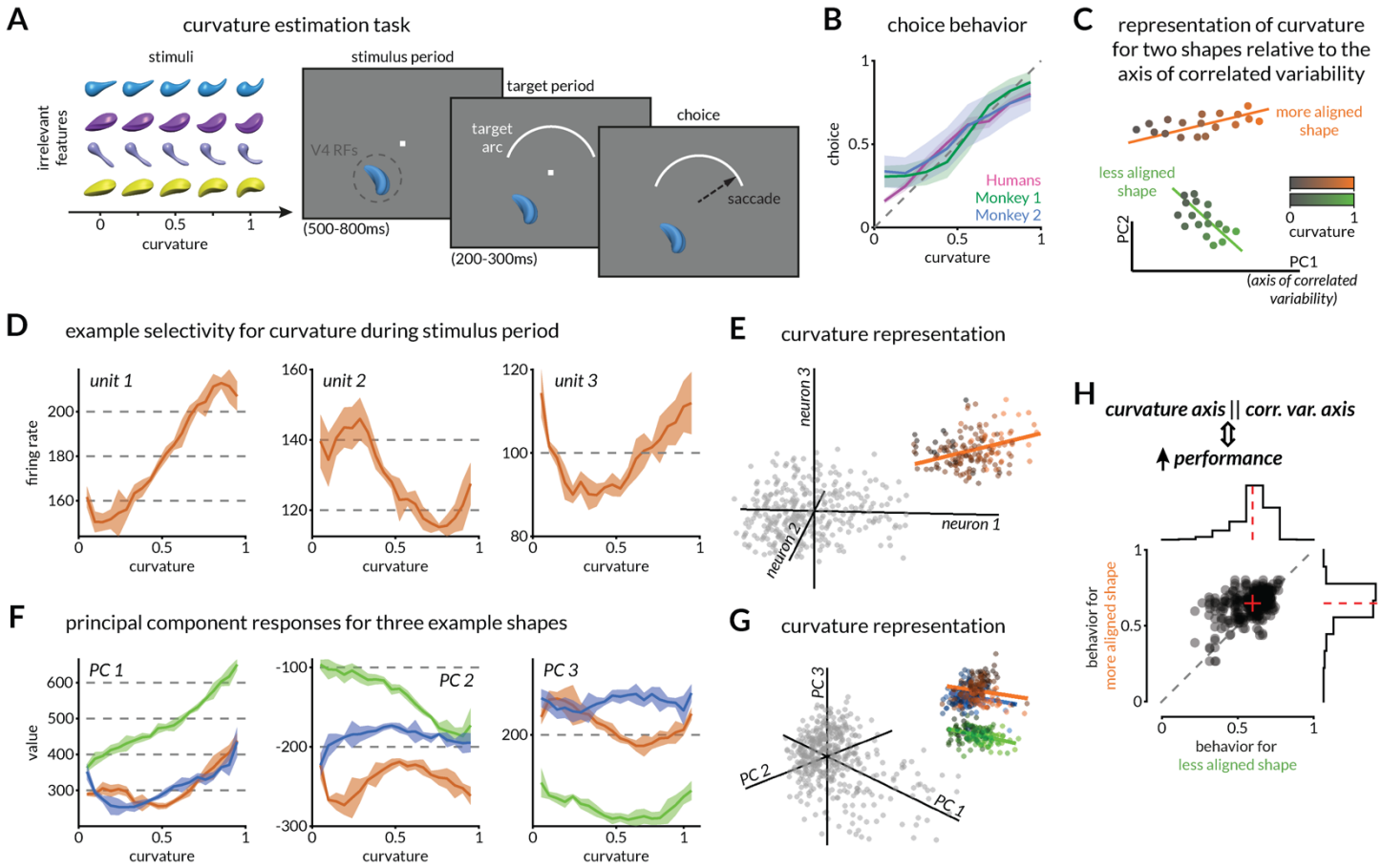

Figure 3: Curvature estimation performance is related to the alignment of the curvature representation with the axis of correlated variability in V4.

A. Schematic of the continuous curvature estimation task. Monkeys report the medial axis curvature of a 3D shape while ignoring irrelevant features like color, orientation, or thickness profile[47]. Stimuli are presented within the joint receptive fields of the recorded V4 neurons. Monkeys are rewarded for making a saccade to the appropriate location on the target arc (left end for straight stimuli, right end for maximally curved stimuli). Monkeys are rewarded in inverse proportion to their estimation error.
B. Behavioral responses of two monkeys (green and blue) plotted as mean reported curvature (normalized from 0 to 1) across various shapes, colors, and orientations. The shaded portion shows the standard deviation. Human participants who indicated the curvature of the same shape stimuli using a slider (pink) performed comparably to the monkeys (r=0.97 between human and monkey choices for matched curvatures).
C. Our hypothesis predicts that performance for a shape with a curvature representation that aligns with the axis of correlated variability is better than one that does not.
D. Example curvature tuning for three V4 multi-units for one shape. Shaded regions indicate standard error of the mean (SEM).
E. Stimulus responses of the units in D plotted relative to each other (black to orange dots are curvature responses to low to high curvatures) and to their respective baseline responses (blue dots).
F. Curvature tuning of the first three principal components across V4 population responses for three shapes (orange, green, and blue). The shaded region represents the standard error of the mean of the eigenvalues for each component.
G. Projections onto the three PCs in F are plotted relative to each other. Conventions as in E.
H. Validation of the prediction in C. Each dot represents performance (defined as average error across all trials) for a pair of shapes tested during the same experimental session. Performance was better for the shape that happened to be better aligned to the axis of correlated variability ($p<0.001$; Wilcoxon signed-rank test). Conventions as in Figure 1C.

175

The axis of correlated variability aligns with planning-related signals in V4 during flexible sensorimotor mapping

Many previous studies have identified pre-motor or, more generally, behavioral planning-related signals even in primarily sensory areas, such as V4[48–54]. These signals could be related to motor efference, drawing attention to context-related signals, or surround modulation and normalization relevant for the task[55–61]. When a behavioral plan is made, downstream areas need to know that plan even more than the sensory signal that guided it. A strong test of our hypothesis that correlated variability aligns with the information that guides behavior is that variability should align more with those planning-related signals than sensory signals. However, in most tasks, including all of those we have discussed so far, the sensory information is not dissociable from motor planning. We therefore designed a variant of our curvature estimation task in which we dissociated sensory from motor by varying the length and angular position of the target arc across trials (Figure 4A). This manipulation required the monkeys to flexibly remap curvature estimates to several different saccade directions (Figure 4B).

We previously showed that V4 population responses reformat during the motor planning period (after arc onset), such that both stimulus curvature and the planned saccade are encoded in V4. Because the planned saccade direction is what most directly guides the upcoming action, we hypothesized that the axis of correlated variability would align more strongly with the saccade than the curvature representation (Figure 4C).

Indeed, both curvature and saccade direction could be decoded from V4 population activity. However, consistent with our hypothesis, when those responses were projected onto the axis of correlated variability, decoding performance for curvature fell more sharply than for saccade prediction (Figure 4D), suggesting that the correlated variability axis was better aligned with the saccade-related than the curvature axis.

This finding is further emphasized by comparing responses to different shapes (Figure 4E). We compared the prediction of our decoder using either the curvature or saccade direction axes with projections onto the correlated variability axis. The difference in curvature prediction accuracy between the dedicated curvature axis and the axis of correlated variability (y-axis) was substantially larger than the difference in saccade prediction accuracy between the dedicated saccade axis and the axis of correlated variability (x-axis). This result indicates that saccade-related representations were more consistently aligned with the axis of correlated variability than curvature representations.

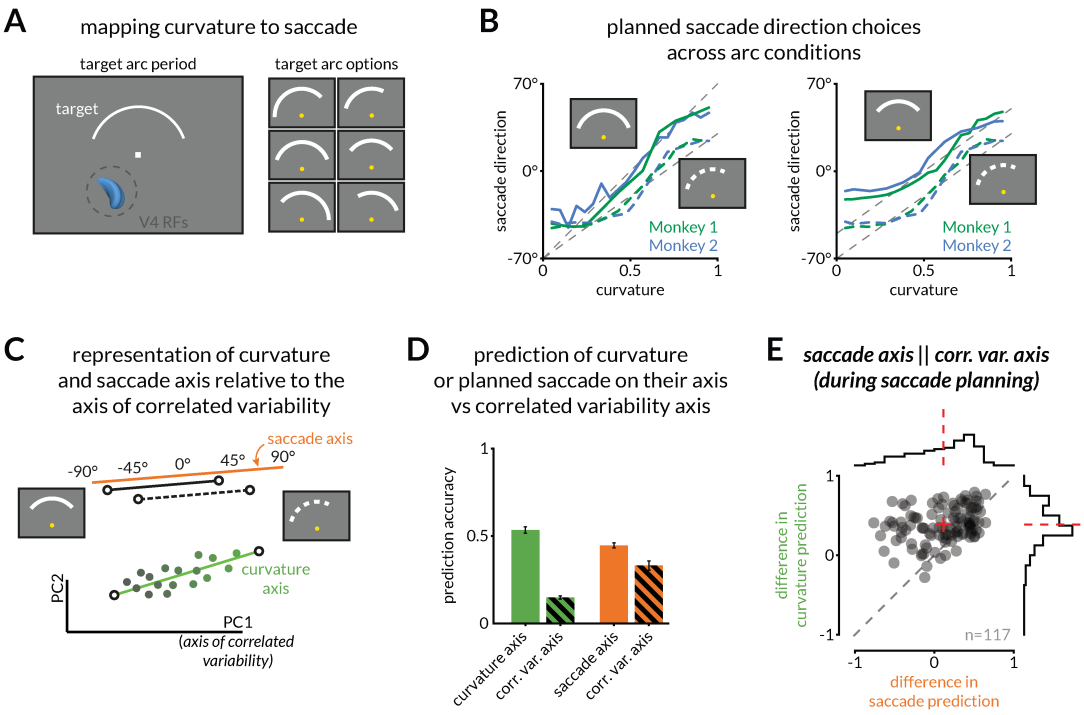

Figure 4: The axis of correlated variability aligns with planning-related signals in V4.

A. In a subset of sessions, the length and angular position of the target arc varied across trials[47]. **B.** Monkeys reconfigure their mapping between curvature judgement and saccade. Mean saccade directions across three arc conditions (the dashed condition is the same across the two panels for comparison) for the two monkeys (colors). The ideal judgement is indicated with the gray dashed line. Left: across two conditions with shared mapping for lower curvatures (-70°) but different mapping for higher curvatures (+70° vs +30°), saccades diverge for high curvatures. Right: across two conditions with the same arc length (100°) but different angular positions (midpoint at 0° vs -20°), psychometric curves for the two conditions have a vertical offset (32 shapes for monkey 1, 34 shapes for monkey 2). **C.** Our hypothesis predicts that if V4 responses are modulated such that they reflect the planned saccade, then the saccade decoding axis should be better aligned with the axis of correlated variability than the curvature decoding axis. **D.** Validation of the prediction in C. The saccade decoding axis is better aligned with the axis of correlated variability than the curvature axis. While both curvature (green) and planned saccade (orange) can be readily decoded from V4 responses, those projected onto the axis of correlated variability are more predictive of saccade (orange dashed) than curvature (green dashed) (mean curvature prediction falls by 0.39 $p<0.001$ t-test, mean saccade direction prediction falls by 0.12 $p<0.001$ t-test; n=117 shapes.) **E.** The difference in curvature prediction accuracy between the curvature axis and the axis of correlated variability (y-axis) is much larger than the difference in the saccade prediction accuracy between the saccade axis and the axis of correlated variability (x-axis) ($p<0.0001$ Wilcoxon signed-rank test; n=117 shapes). Conventions as in Figure 1C.

The axis of correlated variability aligns with the behaviorally relevant visual feature

In the previous experiment, the monkey needed to ignore irrelevant visual features and base decisions on a single, task-relevant feature. Since V4 neurons are selective for multiple visual features and the results above suggest that the axis of correlated variability aligns with behaviorally relevant information (e.g., a motor plan), we next investigated whether this alignment could flexibly shift depending on the sensory feature currently guiding behavior.

To test this, we trained monkeys to perform a two-alternative forced choice task that required decisions based on either curvature or color on randomly interleaved trials (Figure 5A-B). On trials where the two stimuli had the same color, monkeys were rewarded for choosing the more circular (less curved) shape. Conversely, when the two stimuli had the same shape, they were rewarded for making a saccade to the bluer stimulus. Both monkeys successfully made curvature- and color-based choices (Figure 5C), so we divided the data from each session according to the behaviorally relevant feature for each trial. This design allowed us to probe whether the axis of correlated variability in V4 would align more strongly with the representation of curvature on trials in which curvature, rather than color, guided choices (Figure 5D).

For each session, we projected V4 responses onto a common axis of correlated variability, computed from the baseline period before the monkeys knew whether the trial would involve curvature or color discrimination (Figure 5E). When monkeys made curvature-based choices, the projection of neural responses onto the axis of correlated variability reflected the stimulus curvature significantly more accurately than when they made color-based choices. Regardless of the feature that guided choices, the feature axes for color and curvature were orthogonal in V4 representation[47]. This suggests that the axis of correlated variability dynamically aligns with the currently behavior-relevant visual feature.

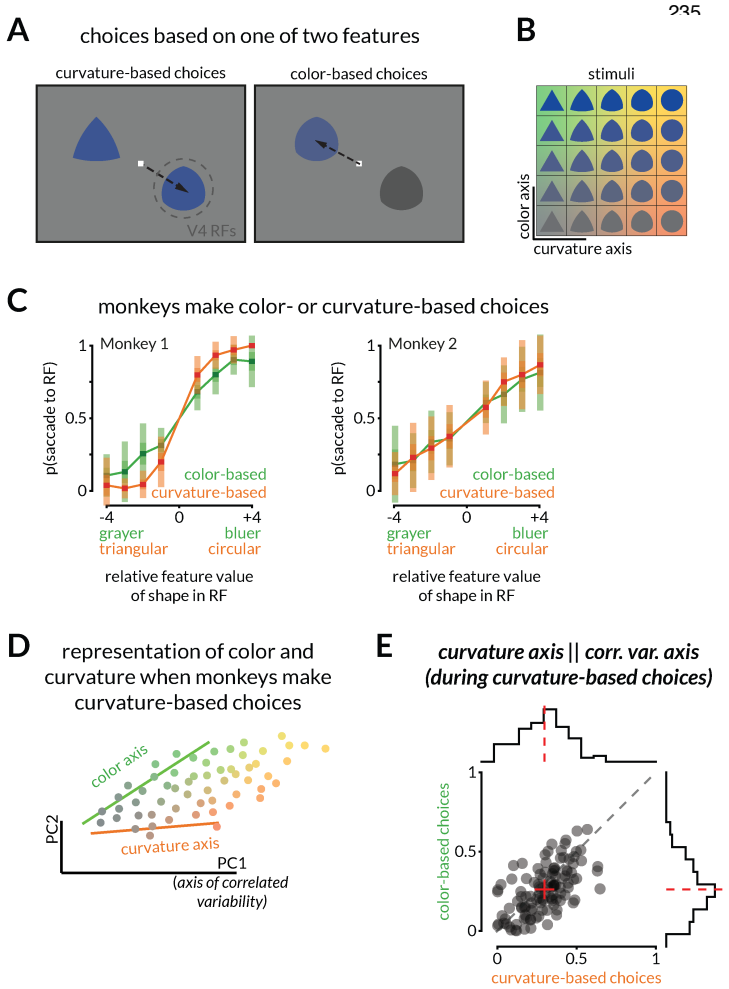

Figure 5: The axis of correlated variability aligns with the visual feature that is relevant for behavior.

A. Schematic of a two-alternative forced choice task in which monkeys make either curvature-based or color-based choices[47]. If the two stimuli have the same color, the monkeys are rewarded if they make a saccade to the more circular shape. If the two stimuli have the same shape, the monkeys are rewarded if they make a saccade to the bluer stimulus. One of the stimuli is presented in the joint receptive fields of V4 neurons.
B. Pairs of stimuli are selected from either the same row (curvature trials) or the same column (color trials).
C. Both monkeys reliably make choices based on color and curvature on interleaved trials (30 sessions for monkey 1, 27 sessions for monkey 2; 268 and 96 color-based trials and 246 and 96 shape-based trials on average for each monkey, respectively).
D. Our hypothesis predicts that the axis in V4 representing the currently relevant feature will be more aligned with the axis of correlated variability.
E. Validation of the prediction in C. When the monkey makes curvature-based choices, the projection of V4 responses onto the axis of correlated variability predicts the curvature of the stimulus more accurately than when the monkey makes color-based choices ($p<0.005$; Wilcoxon signed-rank test). Each dot represents a single session and corresponds to the correlation between stimulus curvature and projection of neural responses onto the axis of correlated variability, averaged across trials when the monkey makes curvature-based choices (x-axis) versus color-based choices (y-axis).

Causal evidence: Behavioral effects of microstimulation are strongest when aligned with the axis of correlated variability

Finally, we causally tested our central hypothesis using electrical microstimulation in the middle temporal area (MT). We chose MT because the effects of microstimulation on the motion judgments have been well-established[59,62–66]. As in our curvature estimation study, we trained monkeys to perform a continuous estimation task. In this case (Figure 6A), they were rewarded for correctly estimating the motion direction of a random dot kinematogram presented within the joint receptive fields of MT neurons recorded on a multielectrode linear probe. On a randomly selected subset of trials, we paired the visual stimulus with microstimulation on one of two electrodes (*long-stim*). The stimulation electrodes were chosen based on the different direction selectivity of the recorded neurons, determined in separate mapping experiments. We quantified the behavioral effect of stimulation as the change in the slope of the psychometric function relating the chosen direction to the motion direction between *long-stim* and *no-stim* conditions (Figure 6B).

We reasoned that a differential behavioral impact of stimulation across two stimulation sites must underlie a difference in how the neural effect of stimulation on those sites aligns with the axis of correlated variability. Following the logic of Moore and Armstrong, 2003[67], we designed a novel technique to measure the neural impact of microstimulation on population responses. In some stimulation trials, we truncated the microstimulation after 50 ms (*short-stim*) and analyzed responses immediately after stimulation (Figure 6C). We chose the stimulation duration for *short-stim* trials such that there was no detectable impact on behavior. We calculated the alignment between the axis of correlated variability and the microstimulation-evoked neural population response change for each electrode on short-stim trials (short-stim unit vectors in Figure 6D).

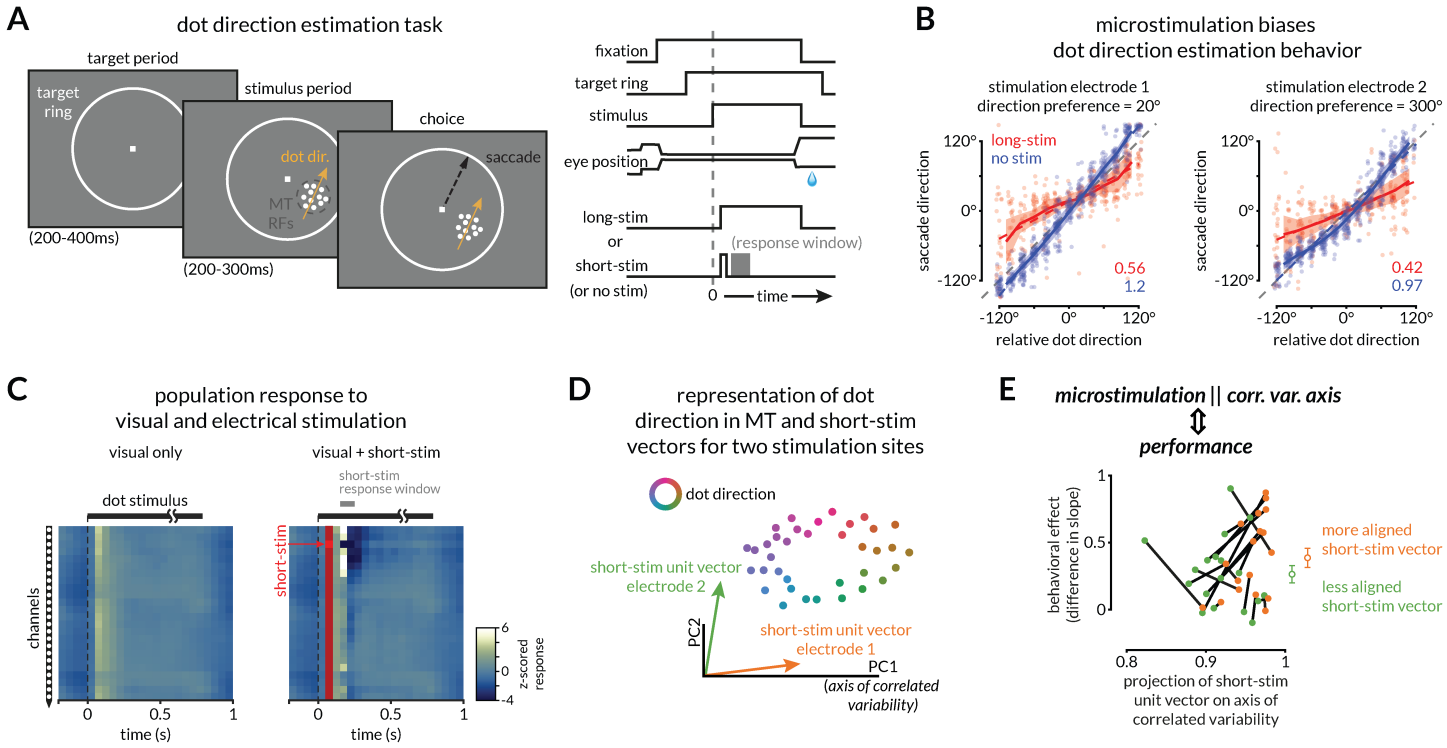

Figure 6: Causal test of our central hypothesis: the behavioral impact of electrical microstimulation is largest when it is aligned with the axis of correlated variability.

A. Schematic of our motion direction estimation task in which monkeys estimate the direction of a random dot kinematogram displayed in the joint receptive fields of MT neurons recorded using linear probes[68]. The trial structure is illustrated on the right. After assessing the direction selectivity of recording sites in an independent mapping experiment, we chose two electrodes for electrical microstimulation. After the monkey fixated a central spot and the target ring appeared, the visual stimulus was presented. Monkeys were rewarded for accurately estimating the direction of motion by making a saccade to a corresponding point on the target ring. On some randomly interleaved trials, either a long period of microstimulation (long-stim; for the full duration of the visual stimulus) or a short microstimulation train (50 ms) was applied on one of two selected electrodes. We analyze the impact of microstimulation on MT responses during the marked analysis period (gray shaded region; 50 to 150 ms after the last stimulation pulse). We compared the monkey's behavior on no-stim and long-stim trials. We selected the short-stim parameters to minimize the behavioral impact of stimulation while maximizing the neural impact.
B. Example direction estimation behavior in no-stim and long-stim trials for one session. The two stimulation electrodes had preferred directions of 20° and 300°, respectively, and stimulation biased the monkey's choices toward those directions. Here, the saccade direction for each trial is plotted against the dot direction relative to the preferred dot direction of the stimulated site. We quantified the behavioral effect as the difference in the slope of the linear fits in the no-stim and long-stim conditions (slopes are indicated in the labels at the bottom right).
C. Example z-scored response histogram across all dot stimuli for the recording sites along the probe for no stimulation (left) and short-stim trials (right). The short 50 ms stimulation pulse train was delivered on the third site (indicated by the red box). We evaluated the effect of the stimulation by using a 100 ms response window after the end of stimulation to calculate a response vector for each of the two stimulation sites in every experimental session.

D. Our hypothesis predicts that choices will be most affected by electrical microstimulation when stimulation moves neural activity along the axis of correlated variability.
E. Validation of the prediction in D. For each session (two connected dots), we identified the stimulation electrode that moved MT population activity in a direction that was more aligned with the axis of correlated variability. For each electrode, we calculated the vector of population activity defined by the firing rate of each neuron on short-stim trials. We then calculated the projection of this vector onto the axis of correlated variability (x-axis; orange dots represent the electrode with the larger projection, so, by definition, orange dots are to the right of their respective green dots representing the other stimulation electrode). The impact of microstimulation on choices (difference in slope between long-stim and no-stim trials) is larger for better-aligned stimulation vectors (a majority of orange dots are above their respective green dots). Mean behavioral effect for more vs less aligned short-stim vector = 0.39 vs 0.26 as depicted by the open circles on the right (error bars indicate SEM).

If choice behavior is preferentially influenced by activity along the axis of correlated variability, then the stimulation site that evokes responses more aligned with this axis should have a stronger impact on the monkeys' choices (Figure 6D). As predicted, the stimulation site that pushed neuronal responses in a direction more aligned with the axis of correlated variability (orange dots in Figure 6E) also had a larger impact on behavior on long-stim trials. This provides causal evidence that neural response fluctuations along the axis of correlated variability are strongly related to behavior.

Together, our results show that the axis of correlated variability does not merely reflect noise to be ignored but instead aligns with—and may amplify (or be a byproduct of response amplification)—the sensory and decision-related signals that guide behavior. This framework reconciles the tight relationship between correlated variability and behavior with the theoretical possibility of decoding information in its presence, pointing to a mechanism by which perception and action can remain flexible, robust, and efficient.

Discussion

Our results demonstrate that shared, trial-by-trial response fluctuations in visual cortex are not simply noise to be averaged out or ignored. Instead, shared variability reflects the information that guides behavior and dynamically tracks changes in task demands. Using complementary correlative, theoretical, and causal approaches and several visually guided tasks, we demonstrated that:

1. *Change detection is most accurate* when stimulus representations vary along the axis of correlated variability;
2. *When feedforward coupling (signal) and recurrent coupling (noise) are aligned in the same circuit*, decoding along the axis of correlated variability is optimal;
3. *Estimates of visual features improve* when the axis of correlated variability aligns with the relevant feature;
4. *The axis of correlated variability flexibly* reflects task demands, from sensory features to action plans;
5. *When perception and action are dissociated*, the axis of correlated variability favors motor intent over sensory evidence;
6. *Causal manipulations are most effective* when aligned with the axis or correlated variability.

Together, these findings support our central hypothesis: the axis of correlated variability reflects the neuronal population activity that is read out to guide behavior. These findings bridge a long-standing gap between theoretical accounts of information-limiting correlations and empirical observations linking them to attention, learning, and behavior. They suggest that variability is not a nuisance, simply tolerated by the brain. Rather, it reflects the substrate of perception and decision-making.

Why measure correlated variability?

In recent years, a staggering number of studies have linked various aspects of flexible, sensory-guided behavior to a remarkably simple measure of shared variability: the mean correlation between the spike count responses of pairs of neurons to repeated presentations of the same stimulus, known as spike count noise correlations (r_{SC}). Correlated variability is modulated by nearly every process that affects perception, including attention, adaptation, learning, task switching, arousal, and stimulus contrast[2,8,9,19,33,69–71], and is linked to behavior on a trial-by-trial basis[20,42]. The prevalence of this relationship suggests a deeper role: that correlated variability provides a window into the neural computations that support cognition, as well as a potential means to restore and enhance them. Specifically, three lines of prior evidence guided our thinking. First, signal and noise correlations are inextricably intertwined. Neurons with similar tuning tend to have higher noise correlations[1,10,16,17,40,42–45], likely due to an underlying shared circuit structure[10,72,73]. Second, we previously showed that in a change detection task, the axis of correlated variability reflects all of the choice-predictive information in V4, even when attention reduces mean correlations. Third, the flexibility of this relationship across stimuli, tasks, and behaviors mirrors the flexible contribution of the visual cortex to visually guided behavior. The responses of neurons in areas like V4 and MT, and their correlated variability, are modulated by attention, task switching, reward, and motor planning[48,55,56,69,74–76], and yet a stable linear decoder can often predict behavior across those contexts[16,23,31,69,77,78]. This suggests that the axis of correlated variability adapts to align with task-relevant dimensions while maintaining a consistent role in behavior.

Reconciling the links between noise, information coding, and performance on perceptual tasks

The link between correlated variability and performance on demanding perceptual tasks is pervasive, but enigmatic. Across species, brain areas, and tasks, decreases in correlated variability have been observed alongside good performance on various perceptual tasks. These modulations are typically interpreted as reflecting an improvement in information encoding: reductions in correlated variability are presumed to reflect less noise, which improves readout[16,23,28]. However, this interpretation is increasingly at odds with theory and data. Theoretical work has demonstrated that optimal decoders can effectively ignore low-dimensional shared variability without compromising performance[16,23,28]. Empirically, we and others have shown that attentional reductions in noise correlations do not improve population-level stimulus encoding[14,30,31].

This raises a paradox: if shared variability does not limit information, why does it change with cognition, and why is it so closely linked to behavior?

Our findings offer a resolution. We propose that the brain does not suppress correlated variability; it uses it. The axis of correlated variability reflects the aspects of population activity that are flexibly read out. Its alignment with task-relevant information explains its links to behavior, attention, and learning, not because it limits information, but because it *reflects what information is used*.

A biologically plausible implementation

Traditionally, models of cortical circuits treat signal and noise as separable quantities. In that scenario, low-dimensional noise can be easily averaged out or ignored so that it is irrelevant to behavior[23,27]. This is contrary to the results presented here, suggesting that, instead, correlated variability is central to the information used to guide behavior.

Our results may simply reflect a biological reality: that signal and noise emerge from the same circuit. The same neurons that encode and transform visual signals have response variability that is bound to be aligned with signal representations[10,72,73]. The axis that accounts for most shared variability may by necessity be the one that reflects the most signal.

To formalize these ideas, we developed a circuit model that links recurrent network dynamics to the structure of shared variability (Figure 2). In the model, if the feedforward and recurrent connectivity are related, then the axis of correlated variability aligns with the feedforward stimulus representation. This alignment *maximizes information transmission by concentrating both signal and correlated noise onto a single axis, which is the optimal configuration for a one-dimensional decoder*, particularly when the signal and noise covary, as is typical in cortex. The model predicts that decoding performance is optimized when the readout axis aligns with the axis of correlated variability. When signal and noise covary, this alignment enhances integration and decoding, aligning with our experimental findings.

Causal manipulations for linking neural population representations to behavior in primates

The primary challenge for using causal manipulations to study the relationship between behavior and correlated variability is the need to measure how manipulations affect neuronal populations. While imaging has been used to measure how electrical stimulation affects neurons in mice[79], it is challenging and uncommon to make these measurements using physiology and in monkeys.

The methods we used for measuring the impact of electrical microstimulation on surrounding populations (Figure 6) are broadly applicable for at least three reasons. First, microstimulation remains a uniquely effective causal manipulation for eliciting behavioral changes, especially in primates. Second, microstimulation leverages the functional organization of the cortex: by varying simple parameters, such as current amplitude, one can easily adjust the number and variety of affected neurons. Finally, microstimulation remains essentially the only causal method for assessing the function of small groups of neurons during human neurosurgery or for prosthetics in humans[80–84]. Understanding the relationship between electrical stimulation, neuronal population activity, and behavior, therefore, has implications both for basic science and translational research. Here, electrical microstimulation in MT (Figure 6) provided a causal test of our hypothesis. Stimulation perturbed population responses along different axes; behavioral impact depended on whether those perturbations aligned with the axis of correlated variability. This result supports the idea that downstream circuits are most sensitive to changes along this axis.

Correlated variability as a handle on cognition

Although correlated variability has primarily been studied from a basic science perspective, our results have translational implications. Correlated variability can be modulated by cognitive states, pharmacological agents, and direct circuit manipulations. If it indeed reflects the information flow from sensory cortex to decision-making circuits, the possibilities for measuring and changing it are endless. Indeed, it is straightforward to modulate correlated variability using existing pharmaceuticals as well as by natural cognitive processes. Our results suggest that correlated variability may provide a powerful biomarker and a potential intervention target for repairing or enhancing perception in health and disease.

In sum, our results reveal that the axis of correlated variability is not just a nuisance byproduct of cortical computation. It is a window into the aspects of neural population activity that flexibly link sensory representations to behavior. By reframing variability as an adaptive feature rather than a limitation, this work offers new insight into the neural basis of perception and points to shared variability as a promising target for understanding and influencing cognition.

Acknowledgements: We are grateful to K. McKracken for providing technical assistance, to John Maunsell for comments on an earlier version of this manuscript, and for helpful comments and suggestions regarding data analysis. We would also like to thank Dr. Lori Holt and Christi Gomez from the Holt lab for guidance and support with the online human psychophysics experiments. This work is supported by Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship (to R.S.), the Simons Foundation (Simons Collaboration on the Global Brain award 542961SPI to M.R.C. and B.D.), and the National Institutes of Health (awards K99EY035362 to R.S., R01EY022930 and RF1NS121913 to M.R.C., R01NS133598 and R01NS137194 to B.D., CRCNS-R01EY034723 to M.R.C. and B.D., and training grant R90DA060338 to Y.X. and B.D.). B.D. benefited from Physics Frontier Center for Living Systems funded by the National Science Foundation (PHY-2317138), and support from the National Institute for Theory and Mathematics in Biology (Simons Foundation award MP-TMPS-00005320 and NSF award DMS-2235451).

Author Contributions:

Study conception: RS, BD, and MRC.
Electrophysiological data collection: Figure 1: AMN and DAR, Figure 3-5: RS, Figure 6: DAR.
Computational and statistical analysis and visualization: RS.
Rate model and related analysis in Figure 2: YX.
Original writing: RS and MRC.
Reviewing and editing: All authors.
Project administration and supervision: BD and MRC.

Conflict of Interest: The authors declare no competing financial interests.

Data and Code Availability: The code used to perform the analyses and generate the figures in this study has been deposited in a public GitHub repository https://github.com/ramanujansrinath/corrvar. Further information and requests for data or custom MATLAB code should be directed to and will be fulfilled by Ramanujan Srinath (ramsrinath@uchicago.edu).

Methods

Experimental Subject Details

We analyzed data from four different datasets, detailed below:

Change detection experiment (Figure 1): Subjects were two adult male rhesus monkeys (*Macaca mulatta*, 8 and 10 kg).

Curvature estimation experiment (Figure 3-4): Subjects were two adult male rhesus monkeys (*Macaca mulatta*, 8 and 9 kg).

Curvature-color 2AFC experiment (Figure 5): Subjects were two adult male rhesus monkeys (*Macaca mulatta*, 11 and 10 kg).

Dot direction estimation experiment (Figure 6): Subjects were two adult male rhesus monkeys (*Macaca mulatta*, 8 and 10 kg).

Before training, all monkeys were surgically fitted with a customized titanium head-post (Crist Instruments Co., Hagerstown, MD). Monkeys were then trained to perform their respective tasks until satisfactory performance was reached. To enable electrophysiological recordings, we chronically implanted monkeys with 96-channel microelectrode arrays (Blackrock Neurotech, Salt Lake City, UT) for recordings in V4 or acute recording chambers over area MT. For the acute recordings, we used 24 or 32-channel linear probes (V- and S-probes; Plexon Inc., Dallas, TX) positioned using grids (Crist Instruments Company Inc., Hagerstown, MD) and advanced using a hydraulic microdrive (Kopf Instruments, Tujunga, CA). All animal procedures for experiments in Figures 1, 3, 4, and 6 were approved by the Institutional Animal Care and Use Committees of the University of Pittsburgh and Carnegie Mellon University, where the electrophysiological and psychophysical data were collected. Animal procedures for experiments in Figure 5 were approved by the Institutional Animal Care and Use Committees of the University of Chicago. Additionally, all data and analyses in this study are reported in accordance with ARRIVE animal use and reporting guidelines.

Curvature estimation experiment (Human version)

We conducted an online psychophysics experiment in which humans reported the curvature of 3D shapes identical to those used in the monkey version shown in Figure 3. These experiments were approved by the Institutional Review Board (IRB) of Carnegie Mellon University. The experiment was conducted on the online psychophysics platform, Gorilla (www.gorilla.sc[85]), and 40 human subjects (ages 19-64; average age 28) were recruited via Prolific (www.prolific.co) between February 7 and 11, 2022. Participants were randomly assigned to one of two groups that each reported the curvatures of five random shapes and their orientation and curvature variations. Because this study was conducted online, we included additional measures to ensure task compliance. We collected data from three questionnaires, which included compliance, demographics, and feedback, per the IRB protocol. Additionally, we detected their monitor size and scaled images accordingly. We confirmed the image size via a standardized credit card size check protocol.

Experimental Methods

Common Experimentation Apparatus

Visual stimuli were displayed on a 24” ViewPixx monitor (1920 × 1080 pixels; 120 Hz refresh rate) or a CRT monitor (1024×768 pixels; 120 Hz refresh rate), both calibrated to linearize intensity, placed 52-60cm away from the monkey. The behavioral experiments (behavioral monitoring, visual display, reward delivery, experimental, and data syncing) were performed using custom MATLAB software and the Psychophysics Toolbox[86]. A square marker on the screen was flashed at the onset of stimuli, which was captured by a photodiode to synchronize the stimulus display with data acquisition. We monitored eye position using an infrared eye tracker (EyeLink 1000 Plus; SR Research). Spiking activity, local field potentials, eye position, and task events were recorded at 30K samples/s using either CerePlex E headstage and CerePlex amplifier (Blackrock Neurotech, Salt Lake City, UT) or Trellis software and Ripple recording hardware (Ripple, Salt Lake City, UT).

Common behavioral, electrophysiological recording, and analysis considerations

Filtering and spike thresholding: We band-pass filtered (250-5000 Hz) the raw electrical activity (acquired at 30K samples/s) and detected threshold crossing timestamps on each recording channel with a manually set threshold (2-3x RMS signal value for each channel). These spiking events, the original raw data, and stimulus-locked photodiode activity were all saved at 30 KHz, and the eye tracking signals were saved at 2 KHz. In this study (as in previous studies from which these data originate), we did not distinguish between sorted single-unit and multi-unit activity.

Baseline response: In all datasets, we calculated the baseline response during the trial epoch after stable fixation and before the onset of the visual stimulus. During this time, the monkey was fixating on a central dot displayed on a gray screen. In most experiments, we varied this duration between 150 and 250 ms, drawing randomly for each trial from a uniform distribution to prevent the monkey from learning the precise timing of the task. We calculated the spike rate for each unit during a fixed, minimum duration for each session.

Stimulus response: We calculated the stimulus-evoked spike rate after a latency of 50 ms (to allow for latency of responses in V4 and MT) during the stimulus display epoch for trials with stable fixation. The details of stimulus durations were specific to each experiment and are detailed below.

Neuron inclusion: For all population-level analyses, we only included units if their average stimulus-evoked response was at least 10% higher than during the baseline, gray-screen period.

Trial inclusion: In all datasets, we analyzed only trials in which the monkey completed either correctly or incorrectly. We excluded trials during which the monkey made a premature saccade to break fixation or those where we detected spurious electrical noise artifacts in the neural recordings.

Calculation of the axis of correlated variability: We performed Principal Components Analysis (PCA) on the baseline activity for each session and defined the axis of correlated variability as the first principal component of this activity. In previous studies[2,12,33], we have calculated this axis, spike count correlations, or the covariance matrix using repeated presentations of the same stimulus. We repeated this analysis for all three datasets and found that the correlated variability axes calculated from baseline and evoked responses were extremely similar (Figure S1). Since the baseline response can be measured during every trial, we opted to calculate the axis of correlated variability using this response.

Change detection experiment

We analyzed data from a previously published dataset in which monkeys performed a cued attention change-detection task while we recorded neuronal activity from area V4[9,31]. Briefly, we analyzed data from a subset of 20 sessions during which the monkeys performed a variant of the change detection task with multiple starting orientations but a constant change amount. During instructional trials, monkeys fixated a central dot while we flashed a single Gabor at the location where the orientation was likely to occur. During subsequent trials, monkeys maintained central fixation while two peripheral Gabor patches were flashed repeatedly (200 ms on, 200-400 ms off). These Gabors had an orientation drawn from a limited set of either [0°, 45°, 90°, 135°] or [0°, 36°, 72°, 108°, 144]. At a random time, the orientation of the Gabor at the cued location changed, and the monkeys were rewarded for making a saccade to the changed stimulus. The orientation change amount was constant throughout the session at either 45° or 36°, depending on the starting orientation set. Spatial attention was manipulated in blocks of trials, each starting with a set of instructional trials. In this study, we only analyzed 80% of the trials in which the change occurred at the cued location. The stimulus responses

were calculated for each flash during the 60-130ms epoch after stimulus onset. Average behavioral performance for each session was calculated as the percentage of hits in detection for each orientation change. The mean hit rate for the most and least aligned orientation change was 66% and 57% respectively. The mean hit rate overall ranged from 29% to 90%. Miscellaneous experimental details of session inclusion, mean firing rates, receptive field mapping, etc., can be found in the original publications[9,21,31]. As in previous studies using this task, we calculated the axis of correlated variability using evoked responses to repeats of the same visual stimulus. This strategy is advantageous because it quantifies correlated variability using the decision formation period. Because our other tasks do not include many repeats of the same stimulus and because it is a stronger test of our hypotheses, for our other data sets, we calculated the axis of correlated variability using baseline responses from the initial period where the monkey fixates a blank screen. We compared axes calculated from stimulus and baseline responses in Figure S1 and discovered that the two are highly similar.

Curvature estimation experiment

We analyzed data from a previously published dataset in which monkeys performed a continuous curvature estimation task while we recorded activity from area V4[47]. Briefly, we analyzed data from a subset of 82 sessions during which monkeys fixated a central dot presented on a gray screen while a randomly generated 3D stimulus was shown in the joint RF of V4 neurons for 550-800ms. For each session, 3-6 stimuli were drawn from a set of 120 base shapes that vary in overall shape (thickness profile, gloss, twist, length, out-of-plane rotation, etc.) or in-plane orientation only or color only. The curvature of the selected shapes was varied, drawn from a uniform distribution across trials in 20 (monkey 1) or 10 (monkey 2) steps. After the stimulus presentation period, a target arc was presented in the upper hemifield. In a majority of sessions, a 140° target arc was presented centrally (82 sessions). In a non-overlapping set of sessions (57 sessions; relevant for analyses in Figure 4), either the angular position (0° or +-20°) or the length of the arc (100° or 140°), or both, were randomized across trials. After the presentation of the arc, the monkeys made a saccade to the arc to indicate their curvature inference. These saccade directions were converted to curvature inference by mapping the possible saccades (-70° to 70° on the 140° centrally presented arc, say) to a scale of 0-1. The reward amount fell linearly along the arc centered on the correct curvature value up to a threshold (+-0.1), after which it fell to 0. We calculated the stimulus-evoked firing rate for each unit during an epoch of 50-550 ms after stimulus onset to allow for V4 response latencies. We calculated the arc-evoked firing rate during an epoch of 0-150 ms after the onset of the arc. Average behavioral performance was calculated for each shape across curvature variations as one minus the average absolute error in curvature judgment. Details about stimulus construction, behavioral timing, reward landscape, and RF mapping can be found in the original manuscript[47].

Curvature-color 2AFC experiment

We analyzed behavioral and neural data from V4 in a two-feature discrimination two-alternative forced-choice (2AFC) experiment. Data from monkey 1 were analyzed and discussed as part of a previous manuscript[47]. We added additional data from monkey 1 and repeated experiments in monkey 2. Briefly, 25 shape stimuli were created by varying the color between gray and blue (isoluminant) in 5 steps and the ‘curvature’ of the stimuli in 5 steps. The curvature was varied by creating homeomorphs of stimuli between an equilateral triangle and a circle using linear interpolation. After the monkey fixated a central dot for 150-250 ms, two shapes that either shared a common curvature value or a common color were sampled from the grid of 25 stimuli and presented in opposite hemifields (with one stimulus location overlapping the joint RFs of V4 neurons). The stimuli were displayed for 200-250 ms, after which the fixation point was removed, serving as a go cue for the monkey to make a saccade to one of the two stimuli. The monkey was rewarded with a drop of juice for selecting the stimulus that was bluer and more circular. We calculated the stimulus response during a window of 50-200 ms

after stimulus onset for each trial. We measured behavioral performance (psychometric) curves by calculating the difference in value of the visual feature (color or curvature) that the monkey can use to guide behavior between the stimulus in the RF and the stimulus in the opposite hemifield. This difference could take eight values between -4 and 4, excluding 0, as we did not present two identical stimuli. Psychometric curves in Figure 5 depict the probability that the monkey chose the stimulus in the RF for each comparison.

Dot direction estimation experiment

Our central hypothesis predicts that external perturbations that are aligned with the axis of correlated variability would have a larger behavioral effect compared to perturbations that are not. We designed a new causal test: we compared the behavioral and neural impacts of electrical microstimulation on two recording sites in the middle temporal area (MT) while monkeys performed a continuous random dot direction estimation task (Figure 6).

Monkeys fixated a central dot on a gray background before a target ring was presented for 200-400 ms. Then, as the monkeys continued to fixate, a dynamic random dot kinematogram was displayed at a location that overlapped the RF of the MT neurons. Monkeys were rewarded for making a saccade to a location on the target ring that corresponded to the direction of the kinematogram, not where the target ring intersects the dot motion vector. Behavioral accuracy was calculated as the slope of the linear relationship between the actual dot direction and the monkey's saccade direction. Unlike the other datasets in this manuscript, the neural data in this experiment were recorded using linear probes with 24 or 32 channels, rather than microelectrode arrays; however, the data were acquired, pre-processed, and analyzed in the same manner. The probes were inserted such that recorded MT units had highly overlapping RFs but different direction tuning preferences. The RFs and direction tuning preferences were measured using an independent experimental protocol to guide stimulus placement and selection of microstimulation sites. In subsets of trials, one of two pre-selected contacts was microstimulated using a biphasic, 200 Hz pulse train with an amplitude ranging between 20 and 40 μA.

Quantifying the behavioral effect of microstimulation: During *long-stim* trials, this pulse train temporally overlapped the visual stimulus. We used these trials to measure the behavioral effect of microstimulation on that channel, calculated as the difference between the slope relating the dot direction to the saccade direction.

Quantifying the neural effect of microstimulation: During *short-stim* trials, the pulse train started ~140 ms after stimulus onset and lasted for 50 ms. We quantified the effect of this short microstimulation train by calculating the stimulation-evoked spike rate for each recording site during the epoch 50 ms after the termination of the last pulse and lasting 100 ms. (We tried 50 ms, 100 ms, and 150 ms epoch durations and found qualitatively similar results.)

Other details of RF mapping, session selection, and behavioral training and timing can be found in a manuscript describing related data[68].

Curvature estimation experiment (Human version)

Since the curvature estimation task is bounded on both ends, subjects routinely overestimate lower curvatures and underestimate higher curvatures. Additionally, we aimed to eliminate the possibility that idiosyncrasies in learning history contributed to any systematic variation in behavioral performance between the two monkeys. It was not feasible to repeat the curvature estimation experiment with many more monkeys, so we designed an online human psychophysics experiment to be run on a large cohort of people, providing us with a baseline to compare monkey performance to. After a set of compliance- and demographics-related questionnaires,

humans performed a slider-based version of the curvature estimation experiment where a horizontal slider was presented with a stimulus image. The initial position of the slider was randomly set across trials, and humans had to use their computer mouse to select a value between 1 and 10 in steps of 0.1. The stimulus images were drawn from the same image set used for the monkeys. We divided our human cohort (n = 40) into two groups. Each group was shown 20 curved variations of 20 shapes (five base shapes at four orientations each). After each trial, the correct curvature was indicated on the slider along with their choice. The maximum time allowed per trial was 4 seconds. We recorded each choice and reaction time. No subject was excluded from analysis.

Network modelling

Figure 2 illustrates a linear-rate network that satisfies the rank-one and alignment assumptions, as well as respecting Dale's law. First, we analyze a simple rank-one circuit whose feed-forward vector and recurrent Perron mode are identical, thereby aligning the noise and stimulus-response directions. Next, we embed an explicit excitatory–inhibitory structure into the same rank-one scaffold, renormalize it so the leading eigenvalue remains, and thereby obtain the Dale-compliant network that produces Figure 2.

Dynamics and general notation

We modelled a linear-rate network of $N = 120$ units. Their activities $\mathbf{r}(t) \in \mathbb{R}^N$ obey

$$\tau \dot{\boldsymbol{r}}(t) = -\boldsymbol{r}(t) + \boldsymbol{G}\boldsymbol{W}_R \boldsymbol{r}(t) + \boldsymbol{G}\boldsymbol{w}_F s(t) + \boldsymbol{\eta},$$

Eq. 1

with $\tau = 1$.

- $s(t) \in \{0,1\}$ is a binary stimulus,
- $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma_\eta^2 \mathbf{I})$ with $\sigma_\eta = 1$,
- $\mathbf{G} = \mathrm{diag}(g_1, \dots, g_N) \succ 0$ are the neuronal gains; here we set $\mathbf{G} = \mathbf{I}$,
- $\mathbf{w}_\mathrm{F}$ is the feed-forward drive,
- $\mathbf{W}_\mathrm{R}$ is the recurrent weight matrix.

Whenever the spectral radius $\rho(\mathbf{W}_\mathrm{R}) < 1$, the steady-state response is Gaussian (the derivation can be found in Appendices A and C of the Supplementary Information):

$$\boldsymbol{r} \mid s \sim \mathcal{N}(\boldsymbol{\mu}_s, \Sigma), \quad \begin{aligned} \boldsymbol{\mu}_s &= (\boldsymbol{I} - \boldsymbol{W}_R)^{-1} \boldsymbol{w}_F\, s, \\ \Sigma &= \sigma_\eta^2 (\boldsymbol{I} - \boldsymbol{W}_R)^{-1} (\boldsymbol{I} - \boldsymbol{W}_R)^{-\top}. \end{aligned}$$

Eq. 2

The signal vector $\Delta\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$ and the linear Fisher ratio

$$J(\mathbf{v}) \;=\; \frac{(\mathbf{v}^\top \Delta\boldsymbol{\mu})^2}{\mathbf{v}^\top \Sigma\, \mathbf{v}}, \qquad \| \mathbf{v} \| = 1,$$

Eq. 3

quantify information along any unit decoder $\mathbf{v}$.

Rank-one symmetric network

We start with the simplest rank-one construction for the feed-forward and recurrent connections.

Feed-forward drive.

Let $N = 120$ be the population size; to simplify the exposition we choose a stimulus acts only on neuron 1 through the unit vector $\mathbf{w} := \mathbf{e}_1 = (1,0,\dots,0)^\top$. In the next section we relax this condition in our Dale's law compliant network used to produce Fig. 2. With a stimulus-amplitude difference of $\beta = 1.0$, we set

$$\mathbf{w}_\mathrm{F} = \beta\, \mathbf{w}.$$

Recurrent weight matrix.

The simplest way to amplify signal and correlated noise along the same axis is to use a rank-one outer product whose Perron eigenvector coincides with $\mathbf{w}$. Choosing a spectral radius of $\rho = 0.4$ gives

$$\mathbf{W}_{\mathrm{R}} = \rho\, \mathbf{w}\mathbf{w}^{\top}, \qquad \mathbf{W}_{\mathrm{R}}\mathbf{w} = \rho\, \mathbf{w},$$

so $\mathbf{w}$ is both the stimulus and the noise direction. This construction satisfies the stimulus–noise alignment condition; Appendix A of the Supplementary Information details the general conditions under which the stimulus-response axis aligns with the dominant noise (slow-mode) axis. Because $\mathbf{W}_{\mathrm{R}}$ is rank 1, the Sherman–Morrison identity yields (see Proposition 1 in Supplementary Methods for the proof):

$$(\mathbf{I} - \mathbf{W}_{\mathrm{R}})^{-1} = \mathbf{I} + \frac{\rho}{1-\rho}\, \mathbf{w}\mathbf{w}^{\top}.$$

With a unit private-noise standard deviation $\sigma_\eta = 1.0$, Eq. 2 becomes (see Proposition 2 in Supplementary Methods for the proof):

$$\Delta\boldsymbol{\mu} = \frac{\beta}{1-\rho}\, \mathbf{w}, \qquad \Sigma = \sigma_\eta^2 \left[\mathbf{I} + \frac{2\rho - \rho^2}{(1-\rho)^2}\, \mathbf{w}\mathbf{w}^{\top}\right].$$

Any unit decoder can be parameterized by an angle θ as $\mathbf{v}(\theta) = \cos\theta\, \mathbf{w} + \sin\theta\, \mathbf{u}_{\perp}$, where $\mathbf{u}_{\perp} \perp \mathbf{w}$ is an arbitrary orthonormal complement. Substituting the expressions above into the Fisher-ratio definition Eq. 3 gives for our symmetric matrix W_{R} the following:

$$J_{\mathrm{sym}}(\theta) = \frac{\beta^2 \cos^2\theta}{\sigma_\eta^2[\cos^2\theta + (1-\rho)^2 \sin^2\theta]},$$

Eq. 4

which reaches its unique maximum at $\theta = 0°$, i.e. when the decoder is aligned with the common stimulus/noise axis $\mathbf{w}$. For general $\mathbf{W}_{\mathrm{R}}$ and $\mathbf{w}_{\mathrm{F}}$ the noise axis is still aligned with the stimulus-response axis, the single-axis decoder remains optimal at $\theta = 0°$ (shown in Appendix B of the Supplementary Information).

Biologically plausible implementation: a Dale-compliant rank-one network

We now redesign the rank-one network framework so that the network respects Dale's law and has a feedforward stimulus structure that better mimics how a tuned population is driven. This is the version that generates the results presented in Fig. 2.

Magnitude profile.

We first place the 120 neurons along a line $x_i = -1 + 2(i-1)/(N-1)$ and assign each neuron:

$$v_i = \exp[-\kappa x_i^2], \qquad \kappa = 4.$$

We normalize the vector as $\mathbf{v} := \mathbf{v}_{\mathrm{mag}} / \| \mathbf{v}_{\mathrm{mag}} \|$ so that $\| \mathbf{v} \| = 1$. This choice makes the recurrent weights largest near the middle and smaller toward the ends, producing a smooth "bump" profile.

E/I sign pattern.

We divide the neurons into excitatory (80 %) and inhibitory (20 %) groups. Rows 1–96 receive a $+$ sign (excitatory) and rows 97–120 a $-$ sign (inhibitory). Element-wise multiplication produces the signed vector

$$\mathbf{z} = \mathrm{sgn} \odot \mathbf{v}, \qquad d := \mathbf{v}^{\top}\mathbf{z} \in (0,1).$$

Because most entries are positive, d is close to 1 ($d \approx 0.985$ with the default parameters). We will divide by d in the outer product below to keep the Perron eigenvalue exactly ρ, ensuring that the Dale's law compliant network inherits the same overall gain and time-scale as the simpler symmetric network.

Recurrent matrix.

Combining magnitude and sign vectors yields a single rank-one matrix

$$\boxed{\mathbf{W}_{\mathrm{R}}^{\mathrm{Dale}} = \frac{\rho}{d}\,\mathbf{z}\,\mathbf{v}^{\top}} \qquad (0 < \rho < 1).$$

Eq. 5

Why the factor $1/d$*?* Right-multiplying by $\mathbf{z}$ gives $\mathbf{W}_{\mathrm{R}}^{\mathrm{Dale}}\mathbf{z} = (\rho/d)\,\mathbf{z}(\mathbf{v}^{\top}\mathbf{z}) = \rho\,\mathbf{z}$. Thus $\mathbf{z}$ is the Perron eigenvector with exactly the same eigenvalue ρ used in the simple symmetric network.

Feed-forward drive.

To mimic the symmetric network we inject the stimulus *along* the Perron vector:

$$\mathbf{w}_{\mathrm{F}} = \beta\,\hat{\mathbf{z}}, \qquad \hat{\mathbf{z}} := \frac{\mathbf{z}}{\|\,\mathbf{z}\,\|} = \mathbf{z} \quad \text{(signs only, norm } = 1).$$

Equivalence to the symmetric rank-one model.

Although it looks different, $\mathbf{W}_{\mathrm{R}}^{\mathrm{Dale}}$ is still rank 1 and shares the same non-zero eigenvalue ρ. Therefore, the inverse used for the symmetric network can be reused, with the replacement $\mathbf{w} \mapsto (\mathbf{z}, \mathbf{v})$.

Using the Sherman–Morrison identity in exactly the same fashion as before, we obtain (see Proposition 3 in Supplementary Methods for the proof):

$$(\mathbf{I} - \mathbf{W}_{\mathrm{R}}^{\mathrm{Dale}})^{-1} = \mathbf{I} + \alpha\,\mathbf{z}\mathbf{v}^{\top}, \qquad \alpha = \frac{\rho}{(1-\rho)d}.$$

Substituting this expression into Eq. 2 gives (see Proposition 4 in Supplementary Methods for the proof):

$$\Delta\boldsymbol{\mu} = \frac{\beta}{1-\rho}\,\hat{\mathbf{z}},$$
$$\Sigma = \sigma_{\eta}^{2}[\mathbf{I} + \alpha(\mathbf{z}\mathbf{v}^{\top} + \mathbf{v}\mathbf{z}^{\top}) + \alpha^{2}\mathbf{z}\mathbf{z}^{\top}].$$

With a decoder $\mathbf{v}(\theta) = \cos\theta\,\hat{\mathbf{z}} + \sin\theta\,\mathbf{u}_{\perp}$, where $\mathbf{u}_{\perp} \perp \hat{\mathbf{z}}$ is any unit vector. Algebra identical to the symmetric case yields

$$J_{\mathrm{Dale}}(\theta) = \frac{\beta^{2}\cos^{2}\theta}{\sigma_{\eta}^{2}(1-\rho)^{2}\,[1 + \kappa\cos^{2}\theta]}, \qquad \kappa = \frac{2\rho}{1-\rho} + \frac{\rho^{2}}{(1-\rho)^{2}d^{2}},$$

Eq. 6

which again peaks at $\theta = 0^{\circ}$. For the default parameters $\rho = 0.4$ and $d \approx 0.985$ we obtain $\kappa \approx 1.79$.

Because both rank and leading eigenvalue are preserved, all subsequent analytic results—noise alignment, decoder optimality, and Fisher-ratio shape—carry over unchanged.

Statistical Analysis and Quantification

Calculation of stimulus axes and comparison with behavior

To quantify the visual information content that is aligned with the axis of correlated variability, we calculate the correlation of the projection of evoked responses on the axis of correlated variability with the stimulus feature. Where appropriate, we also calculate a feature-specific decoder by training a cross-validated linear regression model and compare the decoder performance with the correlation found above. This common analysis across all our datasets forms the scaffolding of the various tests of our central hypothesis. The details of these analyses are experiment-dependent and are detailed below:

Change detection experiments: We first projected the responses for all oriented gratings on the axis of correlated variability. Then, for each starting and change orientation pair, we calculated the performance of a

leave-one-out cross-validated linear model for classifying the two orientations. Of all pairs, we selected the most and least aligned pairs and compared the behavioral performance for those pairs (Figure 1C).

Curvature estimation experiments: For each recording session and base shape, we first projected all stimulus-evoked responses onto the axis of correlated variability. We correlated these projections with the curvature feature value. In other words, we compared the performance of a linear curvature decoder trained on the stimulus-evoked responses projected onto the axis of correlated variability. For each pair of shapes in a given session, we defined the most and least aligned shapes as the ones that had the highest and lowest correlations (or decoding performance). We then compared the average behavioral performance (one minus the average absolute behavioral error) for those two shapes (Figure 3H).

Arc manipulation experiments: To test whether behavioral planning-related activity or feature-related activity varies along the axis of correlated variability, we trained two linear decoders on the activity immediately following the onset of the target arc – one to decode the curvature of the visual stimulus (like above) and one to decode the planned saccade. We compared the performance of these decoders with the performance of curvature and saccade decoders trained on the same responses but projected onto the axis of correlated variability (Figure 4D). We did this separately for each shape (although previous results suggest that a shape-general curvature and saccade decoder would also work just as well[47]). We calculated the difference between the decoder prediction accuracies of the curvature decoders and the saccade decoders. We found a larger drop in curvature decoding performance when projected onto the axis of correlated variability (Figure 4E).

Curvature-color 2AFC experiments: To test whether the information about the visual feature that the monkey uses to guide behavior is the one has the larger projection on the axis of correlated variability, we split each session into two sets of trials – ones in which the monkey uses color to guide choices (trials during which the curvature of the two shapes was the same), and ones in which the monkey uses curvature to guide choices (trials during which the color of the two shapes was the same). We projected the responses of these trials on the axis of correlated variability and correlated them with the respective feature value. We found that when the monkey uses curvature to guide decisions, curvature can be decoded on the axis of correlated variability, but not when the monkey uses color to guide decisions (Figure 5E).

Dot direction estimation experiments: To causally test if stimulating neural activity along the axis of correlated variability would have a greater behavioral effect (versus stimulating orthogonal to it), we quantified the neural and behavioral effects of microstimulation. First, we calculated the projection of the effect of microstimulation on the neural responses (measured as the short stimulation-evoked population response vector; details above) on the axis of correlated variability. We performed this analysis separately for both stimulation sites and identified the site with the larger projection. Then, from the long-stim trials, we calculated the size of the behavioral effect of microstimulation (measured as the difference in the slope of the psychometric curve with and without stimulation; details above). We plotted the size of the behavioral effect against the projection on the noise axis calculated above (Figure 6E).

Statistical tests

In most cases, we used non-parametric tests (Wilcoxon signed-rank test or the Mann-Whitney U test) to test for the hypothesis that two sets of randomly selected samples come from the same bounded or unbounded distribution. The sample sizes are indicated in the figure panel and/or legend.

References

1. Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. Nat. Neurosci. *14*, 811–819. https://doi.org/10.1038/nn.2842.

2. Cohen, M.R., and Maunsell, J.H.R. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. *12*, 1594–1600. https://doi.org/10.1038/nn.2439.

3. Gregoriou, G.G., Rossi, A.F., Ungerleider, L.G., and Desimone, R. (2014). Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. Nat. Neurosci. *17*, 1003–1011. https://doi.org/10.1038/nn.3742.

4. Gu, Y., Liu, S., Fetsch, C.R., Yang, Y., Fok, S., Sunkara, A., DeAngelis, G.C., and Angelaki, D.E. (2011). Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron *71*, 750–761. https://doi.org/10.1016/j.neuron.2011.06.015.

5. Herrero, J.L., Gieselmann, M.A., Sanayei, M., and Thiele, A. (2013). Attention-Induced Variance and Noise Correlation Reduction in Macaque V1 Is Mediated by NMDA Receptors. Neuron *78*, 729–739. https://doi.org/10.1016/j.neuron.2013.03.029.

6. Luo, T.Z., and Maunsell, J.H.R. (2015). Neuronal Modulations in Visual Cortex Are Associated with Only One of Multiple Components of Attention. Neuron *86*, 1182–1188. https://doi.org/10.1016/j.neuron.2015.05.007.

7. Mayo, J.P., and Maunsell, J.H.R. (2016). Graded Neuronal Modulations Related to Visual Spatial Attention. J. Neurosci. *36*, 5353–5361. https://doi.org/10.1523/JNEUROSCI.0192-16.2016.

8. Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron *63*, 879–888. https://doi.org/10.1016/j.neuron.2009.09.013.

9. Ni, A.M., Ruff, D.A., Alberts, J.J., Symmonds, J., and Cohen, M.R. (2018). Learning and attention reveal a general relationship between population activity and behavior. Science *359*, 463–465. https://doi.org/10.1126/science.aao0284.

10. Poort, J., Wilmes, K.A., Blot, A., Chadwick, A., Sahani, M., Clopath, C., Mrsic-Flogel, T.D., Hofer, S.B., and Khan, A.G. (2022). Learning and attention increase visual response selectivity through distinct mechanisms. Neuron *110*, 686-697.e6. https://doi.org/10.1016/j.neuron.2021.11.016.

11. Ruff, D.A., and Cohen, M.R. (2014). Global cognitive factors modulate correlated response variability between V4 neurons. J. Neurosci. *34*, 16408–16416. https://doi.org/10.1523/JNEUROSCI.2750-14.2014.

12. Ruff, D.A., and Cohen, M.R. (2016). Stimulus Dependence of Correlated Variability across Cortical Areas. J. Neurosci. *36*, 7546–7556. https://doi.org/10.1523/JNEUROSCI.0504-16.2016.

13. Ruff, D.A., Alberts, J.J., and Cohen, M.R. (2016). Relating normalization to neuronal populations across cortical areas. J. Neurophysiol. *116*, 1375–1386. https://doi.org/10.1152/jn.00017.2016.

14. Srinath, R., Ruff, D.A., and Cohen, M.R. (2021). Attention improves information flow between neuronal populations without changing the communication subspace. Curr. Biol. *31*, 5299-5313.e4. https://doi.org/10.1016/j.cub.2021.09.076.

15. Zénon, A., and Krauzlis, R.J. (2012). Attention deficits without cortical neuronal deficits. Nature *489*, 434–437. https://doi.org/10.1038/nature11497.

16. Kohn, A., Coen-Cagli, R., Kanitscheider, I., and Pouget, A. (2016). Correlations and neuronal population information. Annu. Rev. Neurosci. *39*, 237–256. https://doi.org/10.1146/annurev-neuro-070815-013851.

17. Nienborg, H., Cohen, M.R., and Cumming, B.G. (2012). Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annu. Rev. Neurosci. *35*, 463–483. https://doi.org/10.1146/annurev-neuro-062111-150403.

18. Nienborg, H., and Cumming, B. (2010). Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? Curr. Opin. Neurobiol. *20*, 376–381. https://doi.org/10.1016/j.conb.2010.05.002.

19. Yan, Y., Rasch, M.J., Chen, M., Xiang, X., Huang, M., Wu, S., and Li, W. (2014). Perceptual training continuously refines neuronal population codes in primary visual cortex. Nat. Neurosci. *17*, 1380–1387. https://doi.org/10.1038/nn.3805.

20. Cohen, M.R., and Maunsell, J.H.R. (2010). A neuronal population measure of attention predicts behavioral performance on individual trials. J. Neurosci. *30*, 15241–15253. https://doi.org/10.1523/JNEUROSCI.2171-10.2010.

21. Ni, A.M., Bowes, B.S., Ruff, D.A., and Cohen, M.R. (2022). Methylphenidate as a causal test of translational and basic neural coding hypotheses. Proc. Natl. Acad. Sci. *119*, e2120529119. https://doi.org/10.1073/pnas.2120529119.

22. Rosenberg, M.D., Zhang, S., Hsu, W.-T., Scheinost, D., Finn, E.S., Shen, X., Constable, R.T., Li, C.-S.R., and Chun, M.M. (2016). Methylphenidate Modulates Functional Network Connectivity to Enhance Attention. J. Neurosci. *36*, 9547–9557. https://doi.org/10.1523/JNEUROSCI.1746-16.2016.

23. Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. Nat. Neurosci. *17*, 1410–1417. https://doi.org/10.1038/nn.3807.

24. Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. Nat. Rev. Neurosci. *7*, 358–366. https://doi.org/10.1038/nrn1888.

25. Ruff, D.A., Ni, A.M., and Cohen, M.R. (2018). Cognition as a Window into Neuronal Population Space. Annu. Rev. Neurosci. *41*, 77–97. https://doi.org/10.1146/annurev-neuro-080317-061936.

26. Pitkow, X., Liu, S., Angelaki, D.E., DeAngelis, G.C., and Pouget, A. (2015). How Can Single Sensory Neurons Predict Behavior? Neuron *87*, 411–423. https://doi.org/10.1016/j.neuron.2015.06.033.

27. Kafashan, M., Jaffe, A.W., Chettih, S.N., Nogueira, R., Arandia-Romero, I., Harvey, C.D., Moreno-Bote, R., and Drugowitsch, J. (2021). Scaling of sensory information in large neural populations shows signatures of information-limiting correlations. Nat. Commun. *12*, 473. https://doi.org/10.1038/s41467-020-20722-y.

28. Kanitscheider, I., Coen-Cagli, R., and Pouget, A. (2015). Origin of information-limiting noise correlations. Proc. Natl. Acad. Sci. *112*, E6973–E6982. https://doi.org/10.1073/pnas.1508738112.

29. Rumyantsev, O.I., Lecoq, J.A., Hernandez, O., Zhang, Y., Savall, J., Chrapkiewicz, R., Li, J., Zeng, H., Ganguli, S., and Schnitzer, M.J. (2020). Fundamental bounds on the fidelity of sensory cortical coding. Nature *580*, 100–105. https://doi.org/10.1038/s41586-020-2130-2.

30. Ruff, D.A., and Cohen, M.R. (2019). Simultaneous multi-area recordings suggest that attention improves performance by reshaping stimulus representations. Nat. Neurosci. *22*, 1669–1676. https://doi.org/10.1038/s41593-019-0477-1.

31. Ni, A.M., Huang, C., Doiron, B., and Cohen, M.R. (2022). A general decoding strategy explains the relationship between behavior and correlated variability. eLife *11*, e67258. https://doi.org/10.7554/eLife.67258.

32. Ni, A.M., Ruff, D.A., Alberts, J.J., Symmonds, J., and Cohen, M.R. (2018). Learning and attention reveal a general relationship between population activity and behavior. Science *359*, 463–465. https://doi.org/10.1126/science.aao0284.

33. Ruff, D.A., and Cohen, M.R. (2014). Attention can either increase or decrease spike count correlations in visual cortex. Nat. Neurosci. *17*, 1591–1597. https://doi.org/10.1038/nn.3835.

34. Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature *370*, 140–143. https://doi.org/10.1038/370140a0.

35. Bair, W., Zohary, E., and Newsome, W.T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. J. Neurosci. Off. J. Soc. Neurosci. *21*, 1676–1697. https://doi.org/10.1523/JNEUROSCI.21-05-01676.2001.

36. Ko, H., Cossell, L., Baragli, C., Antolik, J., Clopath, C., Hofer, S.B., and Mrsic-Flogel, T.D. (2013). The emergence of functional microcircuits in visual cortex. Nature *496*, 96–100. https://doi.org/10.1038/nature12015.

37. Ko, H., Mrsic-Flogel, T.D., and Hofer, S.B. (2014). Emergence of feature-specific connectivity in cortical microcircuits in the absence of visual experience. J. Neurosci. Off. J. Soc. Neurosci. *34*, 9812–9816. https://doi.org/10.1523/JNEUROSCI.0875-14.2014.

38. Khosla, M., Williams, A.H., McDermott, J., and Kanwisher, N. (2024). Privileged representational axes in biological and artificial neural networks. Preprint at bioRxiv, https://doi.org/10.1101/2024.06.20.599957 https://doi.org/10.1101/2024.06.20.599957.

39. Cohen, M.R., and Newsome, W.T. (2008). Context-Dependent Changes in Functional Circuitry in Visual Area MT. Neuron *60*, 162–173. https://doi.org/10.1016/j.neuron.2008.08.007.

40. Denman, D.J., and Contreras, D. (2014). The structure of pairwise correlation in mouse primary visual cortex reveals functional organization in the absence of an orientation map. Cereb. Cortex N. Y. N 1991 *24*, 2707–2720. https://doi.org/10.1093/cercor/bht128.

41. Cumming, B.G., and Nienborg, H. (2016). Feedforward and feedback sources of choice probability in neural population responses. Curr. Opin. Neurobiol. *37*, 126–132. https://doi.org/10.1016/j.conb.2016.01.009.

42. Huang, X., and Lisberger, S.G. (2009). Noise Correlations in Cortical Area MT and Their Potential Impact on Trial-by-Trial Variation in the Direction and Speed of Smooth-Pursuit Eye Movements. J. Neurophysiol. *101*, 3012–3030. https://doi.org/10.1152/jn.00010.2009.

43. Smith, M.A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. J. Neurosci. *28*, 12591–12603. https://doi.org/10.1523/JNEUROSCI.2929-08.2008.

44. Solomon, S.G., and Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. Curr. Biol. CB *24*, R1012-1022. https://doi.org/10.1016/j.cub.2014.09.001.

45. Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. Science *327*, 584–587. https://doi.org/10.1126/science.1179867.

46. Averbeck, B.B., and Lee, D. (2003). Neural Noise and Movement-Related Codes in the Macaque Supplementary Motor Area. J. Neurosci. *23*, 7630–7641. https://doi.org/10.1523/JNEUROSCI.23-20-07630.2003.

47. Srinath, R., Czarnik, M.M., and Cohen, M.R. (2024). Coordinated Response Modulations Enable Flexible Use of Visual Information. Preprint at bioRxiv, https://doi.org/10.1101/2024.07.10.602774 https://doi.org/10.1101/2024.07.10.602774.

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

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

50. Zhang, A., and Zador, A.M. (2023). Neurons in the primary visual cortex of freely moving rats encode both sensory and non-sensory task variables. PLOS Biol. *21*, e3002384. https://doi.org/10.1371/journal.pbio.3002384.

51. Haimerl, C., Ruff, D.A., Cohen, M.R., Savin, C., and Simoncelli, E.P. (2023). Targeted V1 comodulation supports task-adaptive sensory decisions. Nat. Commun. *14*, 7879. https://doi.org/10.1038/s41467-023-43432-7.

52. Steinmetz, N.A., and Moore, T. (2010). Changes in the Response Rate and Response Variability of Area V4 Neurons During the Preparation of Saccadic Eye Movements. J. Neurophysiol. *103*, 1171–1178. https://doi.org/10.1152/jn.00689.2009.

53. Squire, R.F., Noudoost, B., Schafer, R.J., and Moore, T. (2013). Prefrontal contributions to visual selective attention. Annu. Rev. Neurosci. *36*, 451–466. https://doi.org/10.1146/annurev-neuro-062111-150439.

54. Moore, T., Armstrong, K.M., and Fallah, M. (2003). Visuomotor Origins of Covert Spatial Attention. Neuron *40*, 671–683. https://doi.org/10.1016/S0896-6273(03)00716-5.

55. Maunsell, J.H.R. (2015). Neuronal Mechanisms of Visual Attention. Annu. Rev. Vis. Sci. *1*, 373–391. https://doi.org/10.1146/annurev-vision-082114-035431.

56. Treue, S. (2001). Neural correlates of attention in primate visual cortex. Trends Neurosci. *24*, 295–300. https://doi.org/10.1016/S0166-2236(00)01814-2.

57. Reynolds, J.H., and Chelazzi, L. (2004). Attentional Modulation of Visual Processing. Annu. Rev. Neurosci. *27*, 611–647. https://doi.org/10.1146/annurev.neuro.26.041002.131039.

58. Ni, A.M., Ray, S., and Maunsell, J.H.R. (2012). Tuned normalization explains the size of attention modulations. Neuron *73*, 803–813. https://doi.org/10.1016/j.neuron.2012.01.006.

59. Ruff, D.A., and Cohen, M.R. (2017). A normalization model suggests that attention changes the weighting of inputs between visual areas. Proc. Natl. Acad. Sci. *114*, E4085–E4094. https://doi.org/10.1073/pnas.1619857114.

60. Sundberg, K.A., Mitchell, J.F., and Reynolds, J.H. (2009). Spatial Attention Modulates Center-Surround Interactions in Macaque Visual Area V4. Neuron *61*, 952–963. https://doi.org/10.1016/j.neuron.2009.02.023.

61. Flevaris, A.V., and Murray, S.O. (2015). Feature-based attention modulates surround suppression. J. Vis. *15*, 29. https://doi.org/10.1167/15.1.29.

62. Murasugi, C., Salzman, C., and Newsome, W. (1993). Microstimulation in visual area MT: effects of varying pulse amplitude and frequency. J. Neurosci. *13*, 1719–1729. https://doi.org/10.1523/JNEUROSCI.13-04-01719.1993.

63. Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. Nature *346*, 174–177. https://doi.org/10.1038/346174a0.

64. Cohen, M.R., and Newsome, W.T. (2004). What electrical microstimulation has revealed about the neural basis of cognition. Curr. Opin. Neurobiol. *14*, 169–177. https://doi.org/10.1016/j.conb.2004.03.016.

65. Bisley, J.W., Zaksas, D., and Pasternak, T. (2001). Microstimulation of cortical area MT affects performance on a visual working memory task. J. Neurophysiol. *85*, 187–196. https://doi.org/10.1152/jn.2001.85.1.187.

66. Nichols, M.J., and Newsome, W.T. (2002). Middle Temporal Visual Area Microstimulation Influences Veridical Judgments of Motion Direction. J. Neurosci. *22*, 9530–9540. https://doi.org/10.1523/JNEUROSCI.22-21-09530.2002.

67. Moore, T., and Armstrong, K.M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. Nature *421*, 370–373. https://doi.org/10.1038/nature01341.

68. Ruff, D.A., Markman, S.K., Kim, J.Z., and Cohen, M.R. (2025). Linking neural population formatting to function. Preprint, https://doi.org/10.1101/2025.01.03.631242 https://doi.org/10.1101/2025.01.03.631242.

69. Xue, C., Kramer, L.E., and Cohen, M.R. (2022). Dynamic task-belief is an integral part of decision-making. Neuron *110*, 2503-2511.e3. https://doi.org/10.1016/j.neuron.2022.05.010.

70. Ruff, D.A., Xue, C., Kramer, L.E., Baqai, F., and Cohen, M.R. (2020). Low rank mechanisms underlying flexible visual representations. Proc. Natl. Acad. Sci. *117*, 29321–29329. https://doi.org/10.1073/pnas.2005797117.

71. Nandy, A.S., Nassi, J.J., and Reynolds, J.H. (2017). Laminar Organization of Attentional Modulation in Macaque Visual Area V4. Neuron *93*, 235–246. https://doi.org/10.1016/j.neuron.2016.11.029.

72. Chadwick, A., Khan, A.G., Poort, J., Blot, A., Hofer, S.B., Mrsic-Flogel, T.D., and Sahani, M. (2023). Learning shapes cortical dynamics to enhance integration of relevant sensory input. Neuron *111*, 106-120.e10. https://doi.org/10.1016/j.neuron.2022.10.001.

73. Hofer, S.B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., Lein, E., Lesica, N.A., and Mrsic-Flogel, T.D. (2011). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nat. Neurosci. *14*, 1045–1052. https://doi.org/10.1038/nn.2876.

74. Hayden, B., and Gallant, J. (2013). Working Memory and Decision Processes in Visual Area V4. Front. Neurosci. *7*. https://doi.org/10.3389/fnins.2013.00018.

75. Brooks, J.X., and Cullen, K.E. (2019). Predictive Sensing: The Role of Motor Signals in Sensory Processing. Biol. Psychiatry Cogn. Neurosci. Neuroimaging *4*, 842–850. https://doi.org/10.1016/j.bpsc.2019.06.003.

76. Martínez-Trujillo, J.C., and Treue, S. (2002). Attentional Modulation Strength in Cortical Area MT Depends on Stimulus Contrast. Neuron *35*, 365–370. https://doi.org/10.1016/S0896-6273(02)00778-X.

77. Williamson, R.C., Cowley, B.R., Litwin-Kumar, A., Doiron, B., Kohn, A., Smith, M.A., and Yu, B.M. (2016). Scaling Properties of Dimensionality Reduction for Neural Populations and Network Models. PLOS Comput. Biol. *12*, e1005141. https://doi.org/10.1371/journal.pcbi.1005141.

78. Srinath, R., Ni, A.M., Marucci, C., Cohen, M.R., and Brainard, D.H. (2024). Orthogonal neural representations support perceptual judgements of natural stimuli. Preprint at bioRxiv, https://doi.org/10.1101/2024.02.14.580134 https://doi.org/10.1101/2024.02.14.580134.

79. Histed, M.H., Ni, A.M., and Maunsell, J.H.R. (2013). Insights into cortical mechanisms of behavior from microstimulation experiments. Prog. Neurobiol. *103*, 115–130. https://doi.org/10.1016/j.pneurobio.2012.01.006.

80. Szelényi, A., Bello, L., Duffau, H., Fava, E., Feigl, G.C., Galanda, M., Neuloh, G., Signorelli, F., Sala, F., and Workgroup for Intraoperative Management in Low-Grade Glioma Surgery within the European Low-Grade Glioma Network (2010). Intraoperative electrical stimulation in awake craniotomy: methodological aspects of current practice. Neurosurg. Focus *28*, E7. https://doi.org/10.3171/2009.12.FOCUS09237.

81. Flesher, S.N., Collinger, J.L., Foldes, S.T., Weiss, J.M., Downey, J.E., Tyler-Kabara, E.C., Bensmaia, S.J., Schwartz, A.B., Boninger, M.L., and Gaunt, R.A. (2016). Intracortical microstimulation of human somatosensory cortex. Sci. Transl. Med. *8*, 361ra141. https://doi.org/10.1126/scitranslmed.aaf8083.

82. Nowik, K., Langwińska-Wośko, E., Skopiński, P., Nowik, K.E., and Szaflik, J.P. (2020). Bionic eye review – An update. J. Clin. Neurosci. *78*, 8–19. https://doi.org/10.1016/j.jocn.2020.05.041.

83. Farnum, A., and Pelled, G. (2020). New Vision for Visual Prostheses. Front. Neurosci. *14*, 36. https://doi.org/10.3389/fnins.2020.00036.

84. Roelfsema, P.R. (2020). Writing to the Mind's Eye of the Blind. Cell *181*, 758–759. https://doi.org/10.1016/j.cell.2020.03.014.

85. Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N., and Evershed, J.K. (2020). Gorilla in our midst: An online behavioral experiment builder. Behav. Res. Methods *52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x.

86. Brainard, D.H. (1997). The Psychophysics Toolbox. Spat. Vis. *10*, 433–436. https://doi.org/10.1163/156856897X00357.