NEUROSCIENCE

Learning and attention reveal a general relationship between population activity and behavior

A. M. Ni, D. A. Ruff, J. J. Alberts, J. Symmonds, M. R. Cohen*

Prior studies have demonstrated that correlated variability changes with cognitive processes that improve perceptual performance. We tested whether correlated variability covaries with subjects' performance—whether performance improves quickly with attention or slowly with perceptual learning. We found a single, consistent relationship between correlated variability and behavioral performance, regardless of the time frame of correlated variability change. This correlated variability was oriented along the dimensions in population space used by the animal on a trial-by-trial basis to make decisions. That subjects' choices were predicted by specific dimensions that were aligned with the correlated variability axis clarifies long-standing paradoxes about the relationship between shared variability and behavior.

The responses of pairs of neurons to repeated presentations of the same stimulus are typically correlated [quantified as noise correlations, or spike count correlations (r_{SC})] (*1*, *2*). Prior electrophysiological studies have shown that these correlations change with cognitive processes that affect perceptual performance (*2*–*4*). However, theoretical work has suggested that this correlated variability may not affect the information encoded by a neuronal population in a manner that influences a subject's decisions (*5*, *6*).

We therefore measured the relationship between neuronal population activity and performance by studying two processes that both improve visual performance but on very different time scales: attention (*7*) and perceptual learning (*8*). By observing attention and learning in the same behavioral trials and neuronal populations, we identified the dimensions of population activity that matter most for behavior.

We recorded from neuronal populations in V4 (*3*, *4*, *7*–*9*) in two rhesus monkeys with chronically implanted microelectrode arrays (*3*). The monkeys detected changes in the orientation of either of two Gabor stimuli (Fig. 1A): one placed within the receptive fields (RFs) of the recorded neurons and one in the opposite hemifield (Fig. 1B). We measured attention effects within a single session and learning effects across sessions (Fig. 1C).

Attention and perceptual learning improved performance and affected neuronal population responses in similar ways (Fig. 2 and figs. S1 and S2). Both processes were associated with decreases in the mean-normalized trial-to-trial variance (Fano factor) of individual units and the correlated variability between pairs of units (Fig. 2, C, D, J, and K) in response to repeated presentations of the same stimulus (figs. S3 and S4). These variability changes occurred only in the context of the task (variability measured during passive fixation was constant throughout training) (Fig. 2, F, G, M, and N).

Recent theoretical work suggests that only correlated variability along the dimensions in neuronal population space that encode task-relevant stimulus information can limit information coding (*5*, *6*). Determining whether correlated variability lies along these dimensions is experimentally unfeasible because it would require recordings from a very large number of neurons over an even larger number of trials.

Instead, we assessed the importance of attention- and perceptual learning–related changes

Department of Neuroscience and Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, PA 15260, USA.
*Corresponding author. Email: cohenm@pitt.edu

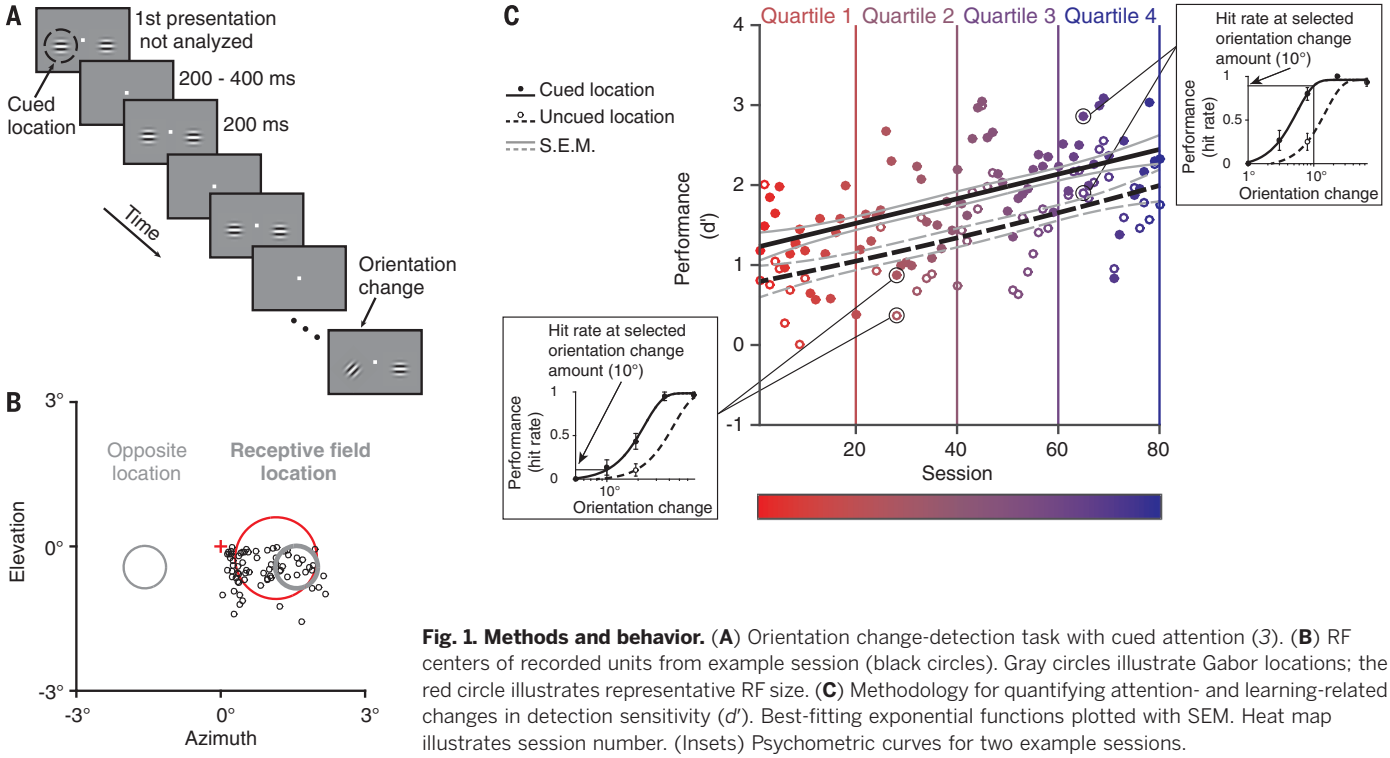


Fig. 1. Methods and behavior. (**A**) Orientation change-detection task with cued attention (*3*). (**B**) RF centers of recorded units from example session (black circles). Gray circles illustrate Gabor locations; the red circle illustrates representative RF size. (**C**) Methodology for quantifying attention- and learning-related changes in detection sensitivity (*d'*). Best-fitting exponential functions plotted with SEM. Heat map illustrates session number. (Insets) Psychometric curves for two example sessions.

in correlated variability by investigating their relationship to behavior. There was a single, robust relationship between correlated variability and perceptual performance, whether changes in performance happened quickly (attention) (Fig. 3, A and B) or slowly (learning) (Fig. 3, C and D). This relationship was robust even when we removed the main effects of attention and learning (Fig. 3, E and F).

We analyzed the responses of V1 neurons (*7*, *8*) in animals performing the same attention task. Unlike in V4, correlated variability in V1 was not correlated with performance (fig. S5).

Both attention and perceptual learning improved the performance of a cross-validated, optimal linear stimulus decoder (fig. S6). However, the relationship between correlated variability in V4 and performance (Fig. 3) seems at odds with theoretical work that suggests most correlated variability should not affect the stimulus information that can be gleaned from an optimal decoder (*6*).

To examine the relationship between correlated variability and performance more directly, we developed a single-trial measure of correlated variability. We performed principal component analysis (PCA) on population responses to the same repeated stimuli used to compute spike count correlations (fig. S3), meaning that the first PC is by definition the axis that explains more of the correlated variability than any other dimension (Fig. 4, A and B, x axis). Consistent with the recent observation that correlated variability is typically low dimensional (*10*–*12*), the variance explained by the first PC was strongly related to the magnitude of correlated variability in each session, even when we accounted for the changes caused by attention and learning (Fig. 4, C and D, and fig. S7) and trial-averaged firing rates (figs. S8 to S11). Like correlated variability (Fig. 3), the proportion of variance explained by the first PC was correlated with behavioral performance (d') across all sessions [Monkey 1, correlation coefficient (R) = −0.42, $P < 10^{-13}$; Monkey 2, R = −0.62, $P < 10^{-15}$].

These analyses show that projection on this first PC is a suitable proxy for pairwise spike count correlations. We used this measure to assess the importance of correlated variability to the monkey by determining whether population activity along this first PC can predict the monkey's choices on a trial-by-trial basis.

Activity along this first PC (and therefore correlated variability) had a much stronger relationship with the monkey's behavior than it would if the monkey used an optimal stimulus decoder. A linear, cross-validated choice decoder (Fig. 4A) could detect differences in hit versus miss trial responses to the changed stimulus from V4 population activity along the first PC alone as well as it could from our full data set (Fig. 4, E and F, and fig. S12). By contrast, although the performance of the stimulus decoder (Fig. 4A) at detecting differences in V4 neuronal population responses to the previous stimulus (the stimulus before the change) versus the changed stimulus was unsurprisingly better overall than

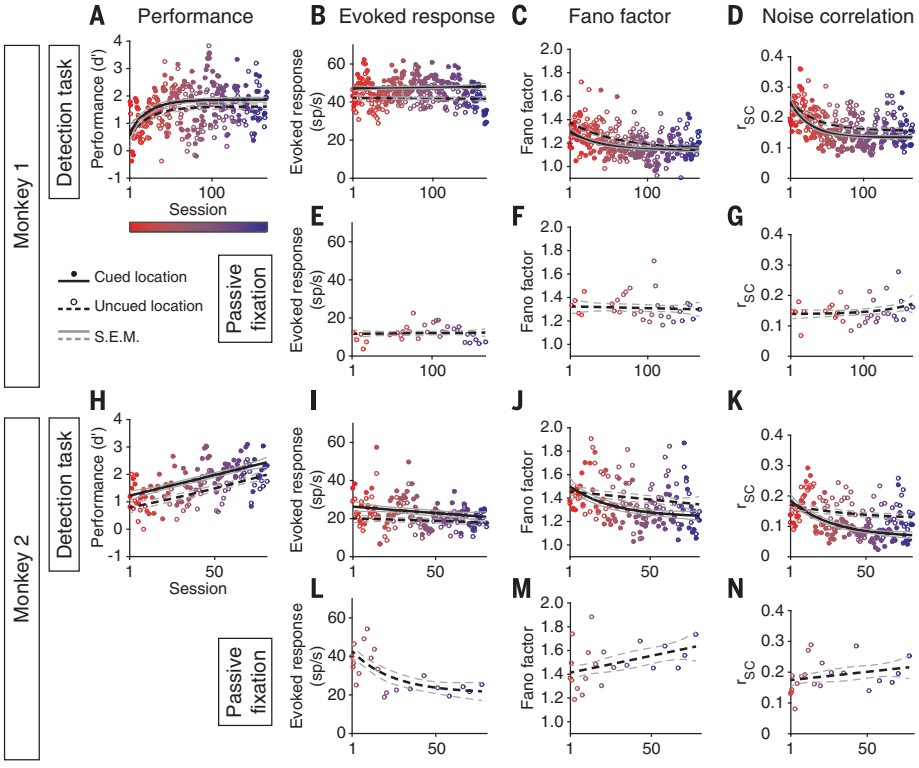


Fig. 2. Summary of behavioral and neuronal effects of attention and perceptual learning. All changes were significantly different than 0 except where indicated (t tests; $P < 10^{-3}$). Conventions are as in Fig. 1C. (**A** and **H**) Sensitivity (d') increased with both attention and learning. (**B**, **E**, **I**, and **L**) Evoked response (firing – baseline rate) increased with attention but did not change consistently with learning or passive fixation ($P > 0.05$). (**C**, **D**, **J**, and **K**) Fano factor and correlated variability decreased with attention and learning, but (**F**, **G**, **M**, and **N**) not during passive fixation ($P > 0.05$).

the performance of the choice decoder (Fig. 4, E and F, insets), the relative influence of the first PC was weaker. The performance of the stimulus decoder was much worse when based on the first PC alone versus our full data set (Fig. 4, E and F).

It is difficult to determine from extracellular recordings whether choice-predictive signals come from a bottom-up, causal relationship between sensory responses and decisions or from trial-to-trial variability from cognitive factors or post-decision signals (*13*). A recent study identifying the directionality of choice-predictive signals in mouse sensory cortex found that they are both bottom-up and top-down in origin (*14*). However, the time course of the choice-predictive activity in our data suggests that it occurs before the decision is made. We based our choice decoder on the first 70 ms of the evoked responses (after accounting for the response latency of V4 neurons). Choice-predictive activity was as strong in the first half of this time frame (60 to 95 ms) as in the second half (96 to 130 ms; paired t test per monkey, $P > 0.05$). That the choice-predictive activity described here was present during the full decision-making period suggests that it did not reflect post-decision feedback.

Our results, combined with functional imaging in humans (*8*) and other multielectrode recording studies (*15*, *16*), suggest that learning is best studied by focusing on populations of neurons. Functional imaging studies, which use measures that are related to the activity of large neural populations, find consistent learning-related changes in both V1 and V4 (*8*, *17*), as opposed to single-unit studies (*8*). Similarly, attention studies suggest that changes in population sensitivity are largely explained by cross-neuron correlations as opposed to single-neuron effects (*3*, *4*).

The robust relationship between correlated variability and perceptual performance suggests that although attention and learning mechanisms act on different time scales (fig. S13), they share a common computation. Some characteristics of this computation are informed by recent studies showing that changes in a low rank modulator can account for the attention-related changes in rate, Fano factor, and correlated variability (*11*, *12*). Attention and learning may decrease the strength of such a modulator by changing the balance of inhibition and excitation (*10*), which may improve information coding and the information that is communicated downstream (*18*).

Our most puzzling finding is that the attention- and learning-related changes in average noise correlation were so closely linked to performance but would likely have a minimal effect on performance if the monkeys read out visual information optimally. Similarly, a prior

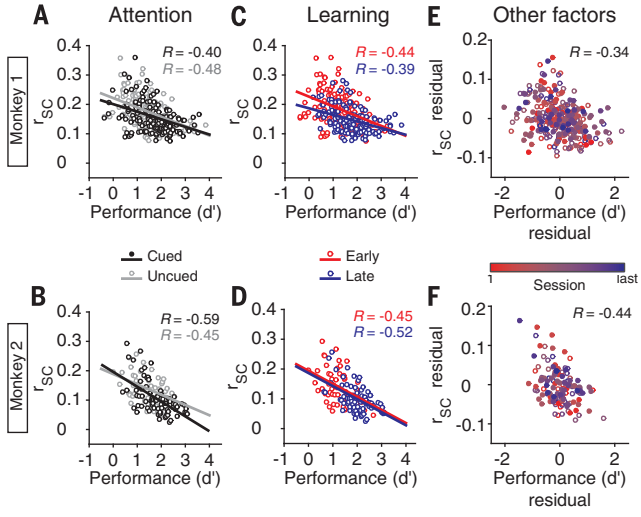


Fig. 3. The relationship between correlated variability and performance is the same for attention and perceptual learning. Mean r_{SC} and d' were significantly correlated across sessions ($P < 10^{-3}$). (**A** and **B**) Relationship between r_{SC} and d' was indistinguishable between attention conditions (Fisher z Pearson-Filon tests; $P > 0.05$). (**C** and **D**) Relationship between r_{SC} and d' was indistinguishable for the first versus second half of learning ($P > 0.05$). (**E** and **F**) Relationships persisted after removing attention and learning effects (residuals of exponential fits in Fig. 2; $P < 10^{-3}$; analyses of variance, $P > 0.05$).

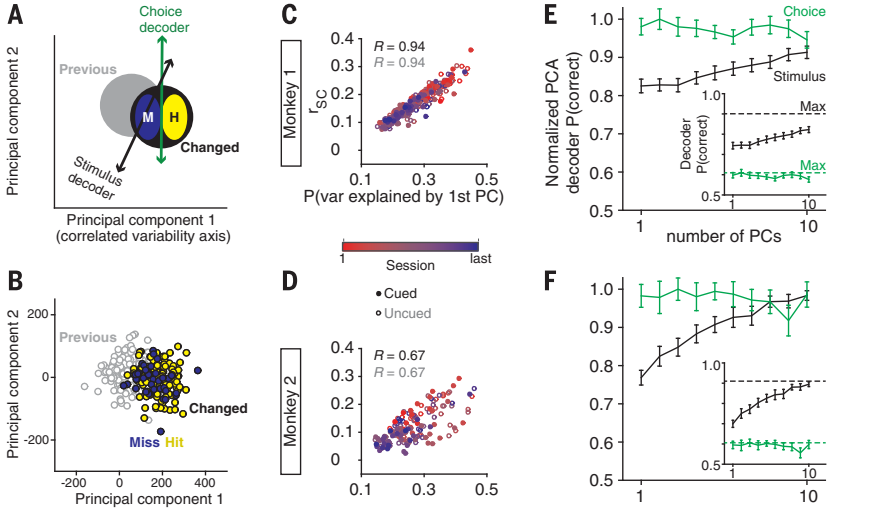


Fig. 4. Correlated variability is more closely aligned with choices than would be expected from an optimal stimulus decoder. (**A**) Schematic showing how we obtained our single-trial measure of correlated variability. We performed PCA on responses to the stimulus before the change (gray) and to the changed stimulus (black). The stimulus decoder detects differences between the responses to the previous and changed stimuli. The choice decoder detects differences between responses to misses and hits. (**B**) Example data set showing that the animal's choices are more aligned with the first PC (x axis) than the difference between the previous and changed stimuli, which depends on both the first and second PC. (**C** and **D**) Mean r_{SC} is highly correlated with proportion of variance explained by the first PC ($P < 10^{-11}$; residuals, $P < 10^{-9}$). (**E** and **F**) Performance of the choice and stimulus decoders normalized to their maximum performance (with the full data set), with SEM. (Inset) Raw decoder performance.

study found that correlations depend on training experience but did not find a relationship between shared variability and information coding (*19*). Correlated variability should only affect the performance of an optimal decoder when it is aligned with the stimulus dimension being decoded (*6*). Therefore, the relationship between correlated variability and performance suggests that our monkeys performed suboptimally.

We thus hypothesize that sensory information is decoded in a way that is optimal for the large number of stimuli and tasks that the animals encounter in their natural environment rather than the particular set of stimuli in our task. Traditionally, optimal decoders are trained to discriminate a particular set of stimuli that vary only in one stimulus dimension. This scenario implies a two-step decision process: identifying the stimulus (to optimize the decoder) and then decoding it. If animals could successfully identify the stimulus, they would perform perfectly on our change-detection task.

Instead, animals may use a more general decoder that could, for example, identify the orientation of any stimulus in any task, meaning that optimal weights would be tuned and noise correlations related to all stimulus features for which the neurons are selective. Noise correlations depend on tuning similarity for all stimulus features (*6*). Therefore, correlated variability is likely aligned with the dimension that is decoded by a general decoder, meaning that noise correlation decreases would improve performance. Several of the studies that suggest monkeys do behave optimally are those that used multisensory stimuli (*20*). Determining whether there is evidence that monkeys use decoders that are optimized for diverse stimuli and tasks will be an important avenue for future work. Our results suggest that the relationship between behavior and population activity is a powerful tool for understanding neural computation.

REFERENCES AND NOTES

1. E. Zohary, M. N. Shadlen, W. T. Newsome, *Nature* **370**, 140–143 (1994).
2. M. R. Cohen, A. Kohn, *Nat. Neurosci.* **14**, 811–819 (2011).
3. M. R. Cohen, J. H. R. Maunsell, *Nat. Neurosci.* **12**, 1594–1600 (2009).
4. J. F. Mitchell, K. A. Sundberg, J. H. Reynolds, *Neuron* **63**, 879–888 (2009).
5. R. Moreno-Bote *et al.*, *Nat. Neurosci.* **17**, 1410–1417 (2014).
6. A. Kohn, R. Coen-Cagli, I. Kanitscheider, A. Pouget, *Annu. Rev. Neurosci.* **39**, 237–256 (2016).
7. J. H. R. Maunsell, *Annu. Rev. Vis. Sci.* **1**, 373–391 (2015).
8. T. Watanabe, Y. Sasaki, *Annu. Rev. Psychol.* **66**, 197–221 (2015).
9. D. A. Ruff, M. R. Cohen, *Nat. Neurosci.* **17**, 1591–1597 (2014).
10. T. Kanashiro, G. K. Ocker, M. R. Cohen, B. Doiron, *eLife* **6**, e23978 (2017).
11. N. C. Rabinowitz, R. L. Goris, M. Cohen, E. P. Simoncelli, *eLife* **4**, e08998 (2015).
12. A. S. Ecker, G. H. Denfield, M. Bethge, A. S. Tolias, *J. Neurosci.* **36**, 1775–1789 (2016).
13. H. Nienborg, B. G. Cumming, *Nature* **459**, 89–92 (2009).
14. S. E. Kwon, H. Yang, G. Minamisawa, D. H. O'Connor, *Nat. Neurosci.* **19**, 1243–1249 (2016).
15. J. M. Jeanne, T. O. Sharpee, T. Q. Gentner, *Neuron* **78**, 352–363 (2013).
16. Y. Yan *et al.*, *Nat. Neurosci.* **17**, 1380–1387 (2014).
17. J. F. Jehee, S. Ling, J. D. Swisher, R. S. van Bergen, F. Tong, *J. Neurosci.* **32**, 16747–53a (2012).
18. D. A. Ruff, M. R. Cohen, *J. Neurosci.* **36**, 7523–7534 (2016).
19. Y. Gu *et al.*, *Neuron* **71**, 750–761 (2011).
20. C. Chandrasekaran, *Curr. Opin. Neurobiol.* **43**, 25–34 (2017).

ACKNOWLEDGMENTS

M.R.C. is supported by U.S. NIH grants 4R00EY020844-03, R01 EY022930, and Core Grant P30 EY008098s; a Whitehall Foundation grant; a Klingenstein-Simons Fellowship; a Sloan Research Fellowship; a McKnight Scholar Award; and a grant from the Simons Foundation. A.M.N. is supported by a fellowship from the Simons Foundation. We thank K. McCracken for technical assistance and J. H. R. Maunsell and A. Kohn for comments on a previous version of this manuscript. A.M.N., D.A.R., and M.R.C. designed the experiments; A.M.N., D.A.R., J.J.A., and J.S. collected the data; A.M.N. performed the analyses; and A.M.N. and M.R.C. wrote the paper. The authors declare no competing financial interests. Data analyzed in this manuscript are available at https://pitt.box.com/v/NiRuffAlbertsSymmondsCohen2017.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6374/463/suppl/DC1
Materials and Methods
Figs. S1 to S13
References (*21*, *22*)

7 June 2017; resubmitted 21 August 2017
Accepted 21 December 2017
10.1126/science.aao0284

Learning and attention reveal a general relationship between population activity and behavior

A. M. Ni, D. A. Ruff, J. J. Alberts, J. Symmonds and M. R. Cohen

Science **359** (6374), 463-465.
DOI: 10.1126/science.aao0284

The neuronal population is the key unit

The responses of pairs of neurons to repeated presentations of the same stimulus are typically correlated, and an identical neuronal population can perform many functions. This suggests that the relevant units of computation are not single neurons but subspaces of the complete population activity. To test this idea, Ni *et al.* measured the relationship between neuronal population activity and performance in monkeys. They investigated attention, which improves perception of attended stimuli, and perceptual learning, which improves perception of well-practiced stimuli. These two processes operate on different time scales and are usually studied using different perceptual tasks. Manipulation of attention and learning in the same behavioral trials and the same neuronal populations revealed the dimensions of population activity that matter most for behavior.

Science, this issue p. 463

Supplementary Material for

Learning and attention reveal a general relationship between population activity and behavior

A. M. Ni, D. A. Ruff, J. J. Alberts, J. Symmonds, M. R. Cohen*

*Corresponding author. Email: cohenm@pitt.edu

Published 26 January 2018, *Science* **359**, 463 (2017)
DOI: 10.1126/science.aao0284

This PDF file includes:

Materials and Methods

The subjects were two adult male rhesus monkeys (*Macaca mulatta*, 8 and 10 kg). All animal procedures were approved by the Institutional Animal Care and Use Committees of the University of Pittsburgh and Carnegie Mellon University.

Experimental Design

The objective of the study design was to measure the behavioral and neuronal effects of both attention and perceptual learning in the same subjects and neuronal populations, using the same task and task trials, such that we could analyze and compare the effects of those two perceptual processes on the activity of simultaneously recorded neurons from visual area V4.

We used an orientation change-detection task for three reasons. First, it allowed us to directly compare the effects of learning (and of attention) on neuronal population activity to those in our previous studies (*3*). Second, because the stimuli preceding the orientation change are identical throughout each recording session, this task provides a very large number of presentations of the same visual stimulus, which is useful for calculating noise correlations and projections onto axes of correlated variability. Finally, these stimuli elicit robust responses in V4 neurons, which have well-characterized tuning curves for orientation.

Behavioral Task

As previously described (*3*), in this orientation-change detection task with cued attention, each monkey fixed their gaze on a small, central spot while two peripheral Gabor stimuli (one overlapping the receptive fields of the recorded neurons, the other in the opposite visual hemifield) synchronously flashed on (for 200 ms) and off (for a randomized period between 200-400 ms) until, at a randomized time, the orientation of one of the stimuli was different from that of preceding stimuli (fig. 1A). The monkey received a liquid reward for making a saccade to the stimulus that changed. Attention was cued in blocks of 125 trials, and a single session consisted of one block of trials with attention cued to the left and one block of trials with attention cued to the right. In each session, the orientation change occurred at the cued stimulus in 80% of trials, and at the uncued (i.e., miscued) stimulus in 20% of trials (all uncued changes used either the middle or largest orientation change).

Before recording, the monkeys were briefly trained on the basic task and the meaning of the attention cue. First, each monkey was trained to report only 90° changes with the attention cue in place. Next, each monkey was trained on the full version of the task, which presented five orientation change amounts (fig. 1C, insets). We began recording after 2-5 days of training on the full version of the task, once the monkey's behavior was stable enough to produce reliable fits of the Weibull function to the psychometric data. The size, location, and spatial frequency of the Gabor stimuli were fixed throughout learning. The orientation of all stimuli before the orientation change was consistent throughout each recording session but changed by 15° between days.

Performance on the task was quantified as detection sensitivity (*d'*), though other behavioral measures gave qualitatively similar results (fig. S1). Sensitivity was calculated for a single orientation change amount per animal (Monkey 1: 29°, Monkey 2: 10°)

throughout all sessions and both attention conditions, based on the Weibull function fit to the psychometric data collected per session.

Recordings

We recorded extracellularly from single units and sorted multiunit clusters (the term "unit" refers to either) in V4 of the left hemisphere using 96-channel microelectrode arrays (Blackrock Microsystems) as previously described (*3*). We presented visual stimuli and tracked eye position as previously described (*9*).

The population size of simultaneously recorded units was 19-42 units (mean 34) per session for Monkey 1 and 6-25 units (mean 15) per session for Monkey 2. The total number of simultaneously recorded unit pairs was 171-861 unit pairs (mean 561) per session for Monkey 1 and 15-300 unit pairs (mean 105) per session for Monkey 2.

The data presented are from 42 d of recording for Monkey 1 and 28 d of recording for Monkey 2. Each day consisted of 1-7 sessions (mean of 3.6/d for Monkey 1; 2.9/d for Monkey 2), for a total of 150 sessions for Monkey 1 and 78 sessions for Monkey 2. Across all sessions, the data presented are from 84,554 unit pairs for Monkey 1 and 8,643 unit pairs for Monkey 2.

Data were collected during passive fixation on 35 d for Monkey 1 and 22 d for Monkey 2.

Statistical Analysis

We based most neuronal analyses on spike count responses between 60-260 ms after stimulus onset (see Decoders section below for exceptions). All analyses used stimulus presentations from correct and miss trials only (i.e., trials in which an orientation change occurred). To minimize the impact of adaptation on our results, we did not analyze the first stimulus presentation in each trial.

We only analyzed a recorded unit if its stimulus-driven firing rate was significantly higher than baseline (Wilcoxon signed rank test; $p < 10^{-10}$). We only included complete sessions, and excluded sessions from analyses if average baseline activity across included units was less than 20 Hz, and outlier sessions were excluded from analyses based on the Tukey method.

We fit sets of data across all sessions with the following exponential equations. For exponential decay of increasing form:

$$y = a(1 - e^{-bx}) + c$$

For exponential decay of decreasing form:

$$y = ae^{-bx} + c$$

For fig. 2, attention effects were quantified with a paired t-test comparing cued vs. uncued trials within each session, and learning effects were quantified during the cued attention condition only with a two-tailed t-test comparing sessions from the first vs. the second half of the total training period, without assuming equal variances.

For fig. 3, we compared the correlation between two variables in the cued attention condition to the correlation between the same two variables in the uncued attention condition using the ZPF test for dependent but non-overlapping Pearson correlation coefficients (*21*). In fig. 3, E and F, we plotted the residuals of the exponential fits for detection sensitivity (*d'*) and correlated variability (r_{SC}) illustrated in fig. 2, A, D, H, and K. To test whether the residuals of the exponential fits for *d'* and r_{SC} contained

attention- or learning-related trends not captured by the exponential fits, we ran an ANOVA per monkey to test the effects of session number and attention condition on the *d'* residual, and an ANOVA per monkey to test the effects of those same two variables on the r_{SC} residual.

Correlated Variability

We defined the correlated variability of each pair of units (quantified as spike count correlation or r_{SC} (*2*)) as the Pearson correlation coefficient between the responses of the two units to repeated presentations of the same stimulus. This measure of r_{SC} represents correlations in noise rather than signal because the visual stimulus was always the same.

To calculate the r_{SC} of one pair of units during one attention condition of one session, we calculated the Pearson correlation between responses to repeated presentations of the stimuli before the orientation change (fig. S3A). These stimuli were the same on every trial in each session (and during all sessions recorded within the same day). We included responses to all stimuli presented prior to the orientation change (except for the first stimulus presentation in each trial, to minimize adaptation effects), during all trials that culminated in an orientation change (see Statistical Analysis section above). We calculated r_{SC} separately for each attention condition. A unit's response to each stimulus was defined as the firing rate of that unit during the 60-260 ms time window after stimulus onset.

Principal Component Analysis

We calculated the principal components for each session of training and for each attention condition, so that we could compare data across sessions of training to analyze perceptual learning effects, and within a single session to analyze attention effects.

For a single session of training, and a single attention condition, we created a matrix of all of the simultaneously recorded units from that session, and all of the included stimuli (the same stimuli used to calculate r_{SC} as described in the Correlated Variability section above) from all of the included trials (see Statistical Analysis section above) recorded in that attention condition that session. We then ran principal component analysis (PCA) on this data matrix. All of the units were included into the PCA, and the PCA pulled the principal components of the entire data matrix.

Decoders

The optimal stimulus decoder (fig. S6) was a linear classifier with leave-one-out cross-validation that was trained to discriminate the responses of the neuronal population to the stimulus before the change (the previous stimulus) from the responses of the neuronal population to the changed stimulus. We quantified the performance of the decoder as the proportion of leave-one-out trials that the decoder discriminated correctly. We measured decoder performance as a function of population size, from a single recorded unit through the maximum number of simultaneously recorded units. The maximum number of simultaneously recorded units per monkey was based on classifier constraints on the pooled covariance matrix (Monkey 1: 30 units; Monkey 2: 10 units). We randomly selected subsets of simultaneously recorded units without replacement 1000 times for each population size. To maximize the number of behavioral trials, we

analyzed all trials in a given day together, focusing only on trials that presented the middle orientation change amount, for which we had collected trials with both cued and uncued orientation changes. Because the middle orientation change amount varied across recording days, we matched the distributions of orientation change amounts across learning in all analyses (after mean matching, n = 37 d for Monkey 1, n = 10 d for Monkey 2).

To avoid artifacts in neuronal firing rates due to eye movements in response to the changed stimulus, we performed decoder analysis on the changed and previous stimulus responses with an abbreviated time window: spike count stimulus responses were measured between 60-130 ms after stimulus onset.

The PCA stimulus decoder (Stimulus decoder; fig. 4A), like the optimal stimulus decoder, was a linear classifier with leave-one-out cross-validation that was trained to discriminate the neuronal population activity in response to the previous stimulus from the neuronal population activity in response to the changed stimulus. However, while we measured the performance of the optimal stimulus decoder as a function of population size, we measured the performance of the PCA stimulus decoder as a function of number of PCs (from the 1st PC only to the maximum testable number of PCs), to test whether the performance of the PCA stimulus decoder improved with increasing number of included PCs.

To calculate the PCs of the neuronal population activity in response to the previous stimulus as well as the PCs of the neuronal population activity in response to the changed stimulus, we first calculated the PCs of the neuronal population activity in response to the previous stimulus (as described in the Principal Component Analysis section above). The neuronal population activity in response to the previous stimuli and the neuronal population activity in response to the changed stimuli were both projected onto those PCs, such that the responses to both the previous and changed stimuli were projected onto identical PCs.

The PCA choice decoder (Choice decoder; fig. 4A) was a linear classifier with leave-one-out cross-validation that was trained to discriminate the neuronal population activity in response to correct choices (hits; when the monkey correctly responded to a changed stimulus) from the neuronal population activity in response to incorrect choices (misses; when the monkey did not respond to a changed stimulus). Thus, the PCA choice decoder discriminated neuronal population activity based on the monkey's trial-by-trial choices, with decoder performance quantified as the proportion of leave-one-out trials for which the decoder correctly discriminated the monkey's choice.

Like with the PCA stimulus decoder, we projected the neuronal population activity in response to correct choices as well as the neuronal population activity in response to incorrect choices onto the same PCs, the PCs of the neuronal population activity in response to the previous stimulus. Thus, all neuronal activity, whether for the PCA stimulus decoder or for the PCA choice decoder, was projected onto identical PCs.

Fig. S1

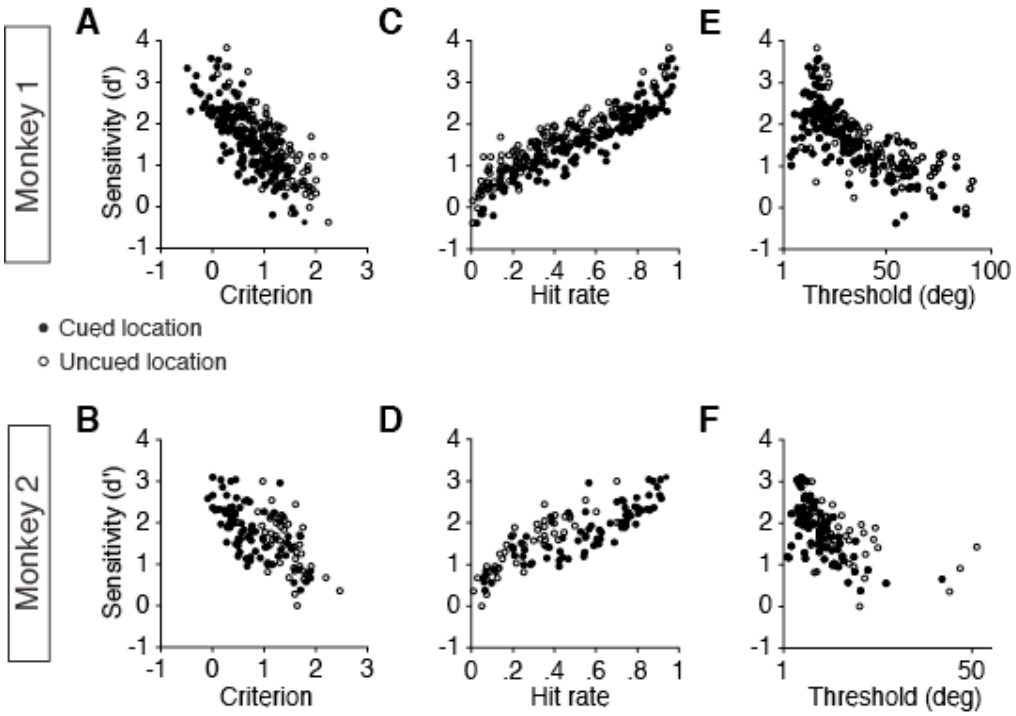


Fig. S1. All behavioral measures were correlated. Though the perceptual performance of each monkey was represented in this study by the monkey's sensitivity (*d'*), other behavioral measures gave qualitatively similar results to sensitivity. Each point represents one attention condition (cued or uncued) for a single session. (**A**) Sensitivity (*d'*) vs. criterion for Monkey 1. Pearson correlation coefficients: cued: $R = -0.78$, $p < 10^{-32}$; uncued: $R = -0.80$, $p < 10^{-31}$. (**B**) Same for Monkey 2. Cued: $R = -0.64$, $p < 10^{-10}$; uncued: $R = -0.64$, $p < 10^{-10}$. (**C**) Sensitivity vs. hit rate for Monkey 1. Cued: $R = 0.92$, $p < 10^{-62}$; uncued: $R = 0.93$, $p < 10^{-59}$. (**D**) Same for Monkey 2. Cued: $R = 0.85$, $p < 10^{-22}$; uncued: $R = 0.88$, $p < 10^{-17}$. (**E**) Sensitivity vs. threshold for Monkey 1. Cued: $R = -0.75$, $p < 10^{-27}$; uncued: $R = -0.70$, $p < 10^{-18}$. (**F**) Same for Monkey 2. Cued: $R = -0.65$, $p < 10^{-10}$; uncued: $R = -0.45$, $p < 10^{-3}$. Number of sessions: Monkey 1: $n = 150$, Monkey 2: $n = 78$.

Fig. S2

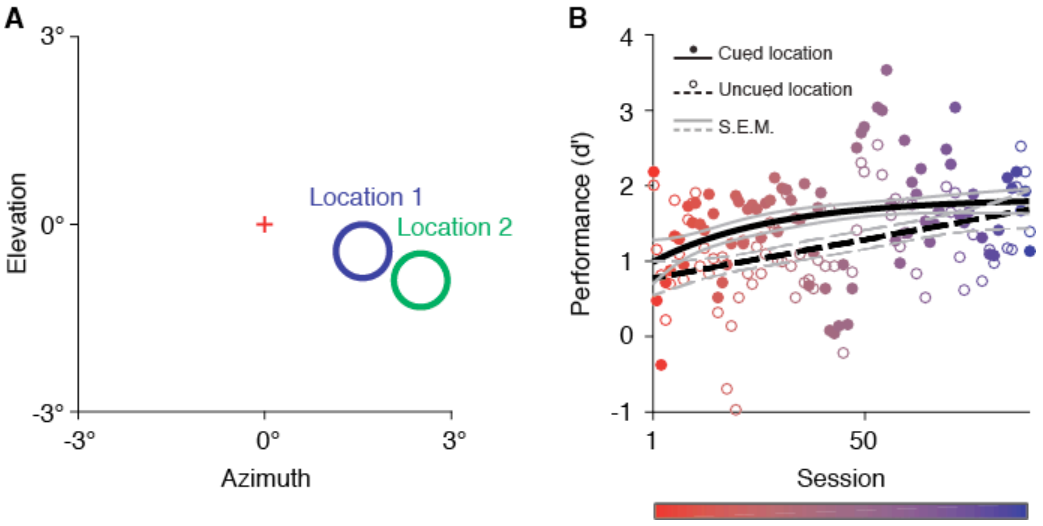


Fig. S2. Behavioral improvement was specific to the trained location. For one subject (Monkey 1), we demonstrated that, as in traditional perceptual learning paradigms, the learning-related improvement in perceptual sensitivity (*d'*) was specific to the trained location. (**A**) Monkey 1 was trained at location 1 first (the location that overlapped the receptive fields of the recorded units; *blue circle*), then at location 2, a nearby location (*green circle*). The monkey fixated a point at the center of the screen (*red cross*). (**B**) As with location 1, the monkey required training to reach a high level of sensitivity (*d'*) at location 2. Plot follows format of fig. 1C. Sensitivity (*d'*) increased with both attention ($p < 10^{-5}$) and learning ($p < 10^{-3}$). Number of sessions = 82.

Fig. S3

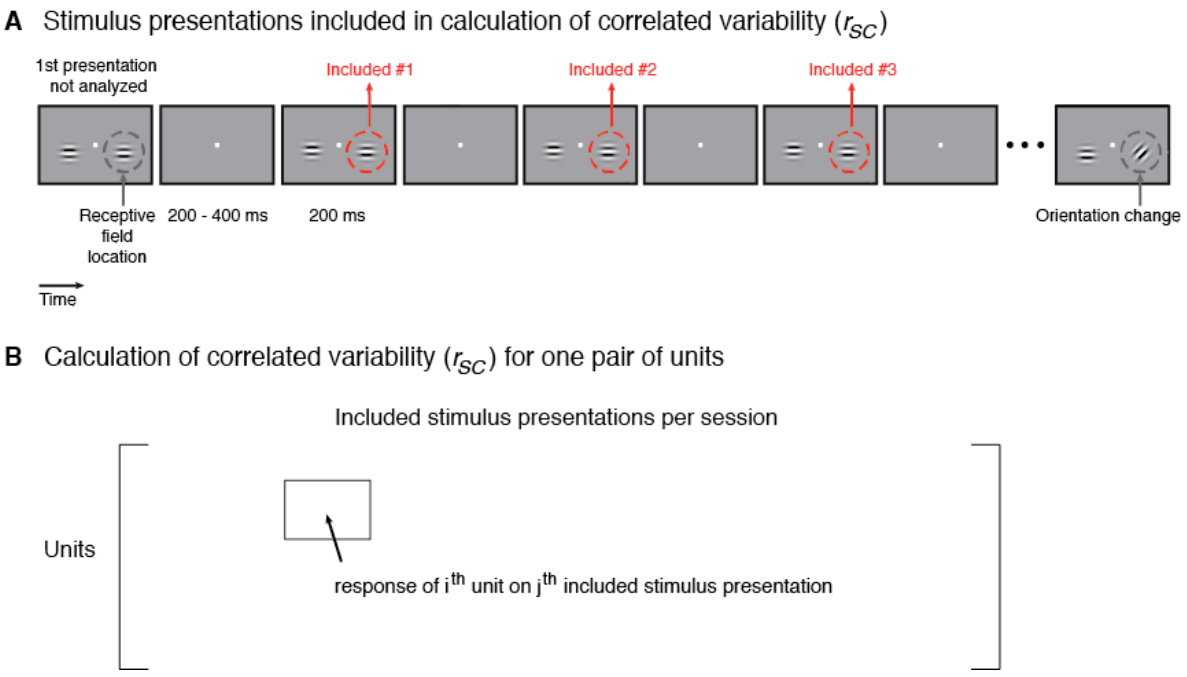


Fig. S3. Calculation of correlated variability (r_{SC}). (**A**) The correlated variability (noise correlation, or, spike count correlation (r_{SC})) of each pair of units was calculated based on spike count responses to repeated presentations of the stimuli before the orientation change (which were the same on every trial in a given session). The first stimulus presentation of each trial was excluded to minimize adaptation effects. Only the stimuli from correct response and missed response trials (i.e., trials in which an orientation change occurred) were included in the calculation of r_{SC}. (**B**) The correlation matrix is based on the number of units by number of stimulus presentations matrix of spike count responses to all included stimulus presentations. The r_{SC} for each pair is the entry of the resulting correlation matrix that reflects the Pearson's correlation coefficient between the responses of the two units. For most analyses (unless otherwise noted), we report the average r_{SC} across all simultaneously recorded pairs in each session.

Fig. S4

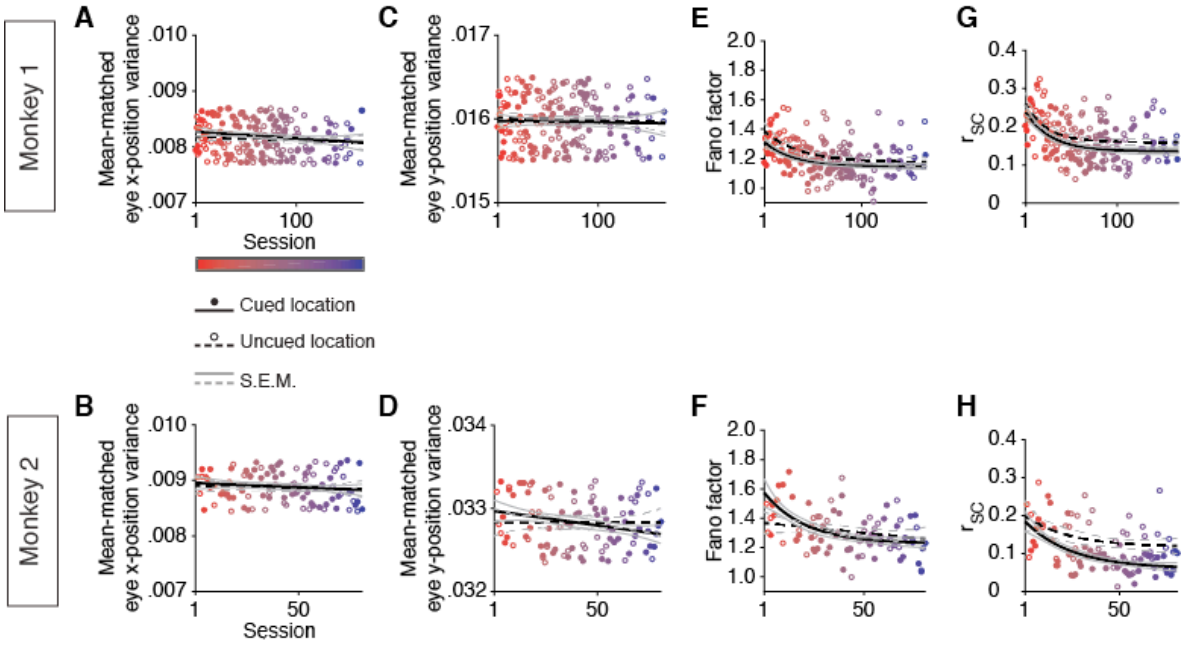


Fig. S4. Decreases in neuronal variability were not due to changes in eye movement variability. In theory, the decreases in Fano factor and correlated variability attributed to attention and perceptual learning in fig. 2 could be caused by attention- and perceptual learning-related differences in eye movement variance (though attention-related differences would be due to differences in eye movement variance between when the monkey attended to the stimulus in the receptive fields of the recorded neurons and when it attended the other stimulus). To control for this possibility, we matched the distributions of eye position variance across attention conditions and sessions. We found that the attention- and learning-related changes in response variability were present even in these subsets of data with matched eye position variance. Each plot follows the format of fig. 1C. Attention and perceptual learning effects quantified as per fig. 2. (**A**) Eye x-position variance was matched across attention conditions ($p = 0.72$) and across sessions to match across learning ($p = 0.20$) for Monkey 1. (**B**) Same for Monkey 2 (attention: $p = 0.40$; learning: $p = 0.17$). (**C**) Same for eye y-position variance for Monkey 1 (attention: $p = 0.30$; learning: $p = 0.64$). (**D**) Same for Monkey 2 (attention: $p = 0.50$; learning: $p = 0.07$). (**E**) Fano factor decreased with both attention ($p < 0.05$) and learning ($p < 10^{-3}$) for Monkey 1. (**F**) Same for Monkey 2 (attention: $p < 0.05$; learning: $p < 10^{-3}$). (**G**) Correlated variability decreased with both attention ($p < 0.05$) and learning ($p < 0.05$) for Monkey 1. (**H**) Same for Monkey 2 (attention: $p < 10^{-4}$; learning: $p < 10^{-4}$). Number of sessions: Monkey 1: $n = 150$, Monkey 2: $n = 78$.

Fig. S5

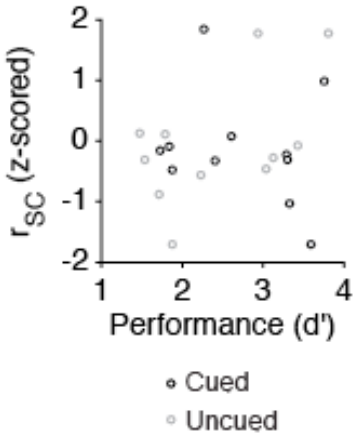


Fig. S5. The relationship between correlated variability and performance is weak in V1. For groups of V1 neurons in two additional animals that were previously well trained to perform the attention task used in the current study (*18*), correlated variability (r_{SC}) was not correlated with performance (*d'*): Pearson correlation coefficient: cued: $R = 0.04$, $p = 0.89$; uncued: $R = 0.53$, $p = 0.07$. Data z-scored within monkey. While differences between the monkeys used in these two studies and small differences between the V1 and V4 tasks (e.g., the V1 task included stimuli of lower contrast, and the attended and unattended stimuli were not as far apart) make a quantitative comparison between the V1 and V4 results difficult, our data suggest that the relationship between r_{SC} and *d'* is stronger in V4 than in V1.

Fig. S6

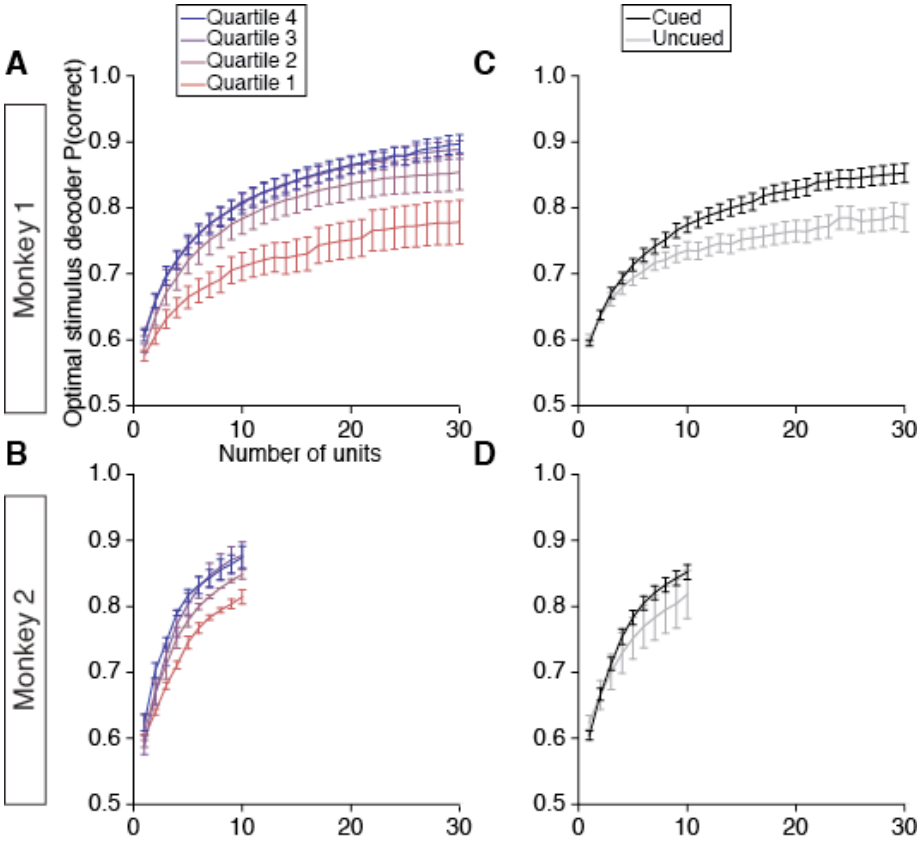


Fig. S6. Optimal stimulus decoder analysis. Using simultaneously recorded units (which are almost certainly a small subset of the units the monkey uses to solve the task), the ability of an optimal, linear, leave-one-out cross-validated decoder to detect changes in the visual stimulus improved with perceptual learning and attention. Decoder performance (y-axis) was measured as a function of population size (x-axis), with maximum pool size based on classifier constraints on the pooled covariance matrix (Monkey 1: 30 units; Monkey 2: 10 units). We randomly selected subsets of units without replacement 1000 times for each population size. We determined the ability of the decoder to discriminate between neuronal responses to the stimulus prior to the orientation change and neuronal responses to the changed stimulus. To avoid artifacts in neuronal firing rates due to eye movements in response to the changed stimulus, we performed decoder analysis on the changed and previous stimuli based on neuronal responses during an abbreviated time window: spike count stimulus responses were measured between 60-130 ms after stimulus onset. We only included trials that presented the middle orientation change amount, for which we had cued and uncued orientation changes, and matched the distributions of middle orientation change amounts across all learning and attention conditions. (**A**) Optimal stimulus decoder performance improved throughout perceptual learning over a long time period (see fig. 1C for learning quartile illustration) for Monkey 1, (**B**) and Monkey 2, (**C**) as well as with attention within each day for Monkey 1, (**D**) and Monkey 2. Error bars are S.E.M. Number of days: Monkey 1: $n = 37$, Monkey 2: $n = 10$.

Fig. S7

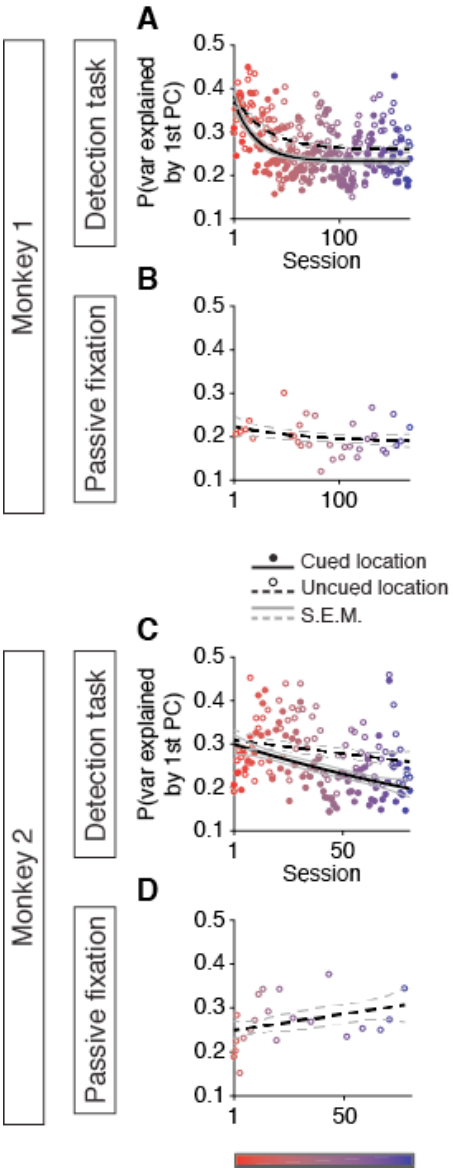


Fig. S7. Both attention and perceptual learning are associated with decreases in the proportion of the total variance explained by the first principal component (PC). Each plot follows format of fig. 1C. Attention and learning effects quantified as per fig. 2. (**A**, **C**) Effects of attention (Monkey 1: $p < 10^{-9}$; Monkey 2: $p < 10^{-8}$) and learning (Monkey 1: $p < 10^{-3}$; Monkey 2: $p < 10^{-7}$) during the change-detection task. (**B**, **D**) These effects were not observed during passive fixation (Monkey 1: $p = 0.12$; Monkey 2: $p = 0.30$). Number of sessions: detection task: Monkey 1: $n = 150$, Monkey 2: $n = 78$; passive fixation: Monkey 1: $n = 35$, Monkey 2: $n = 22$.

Fig. S8

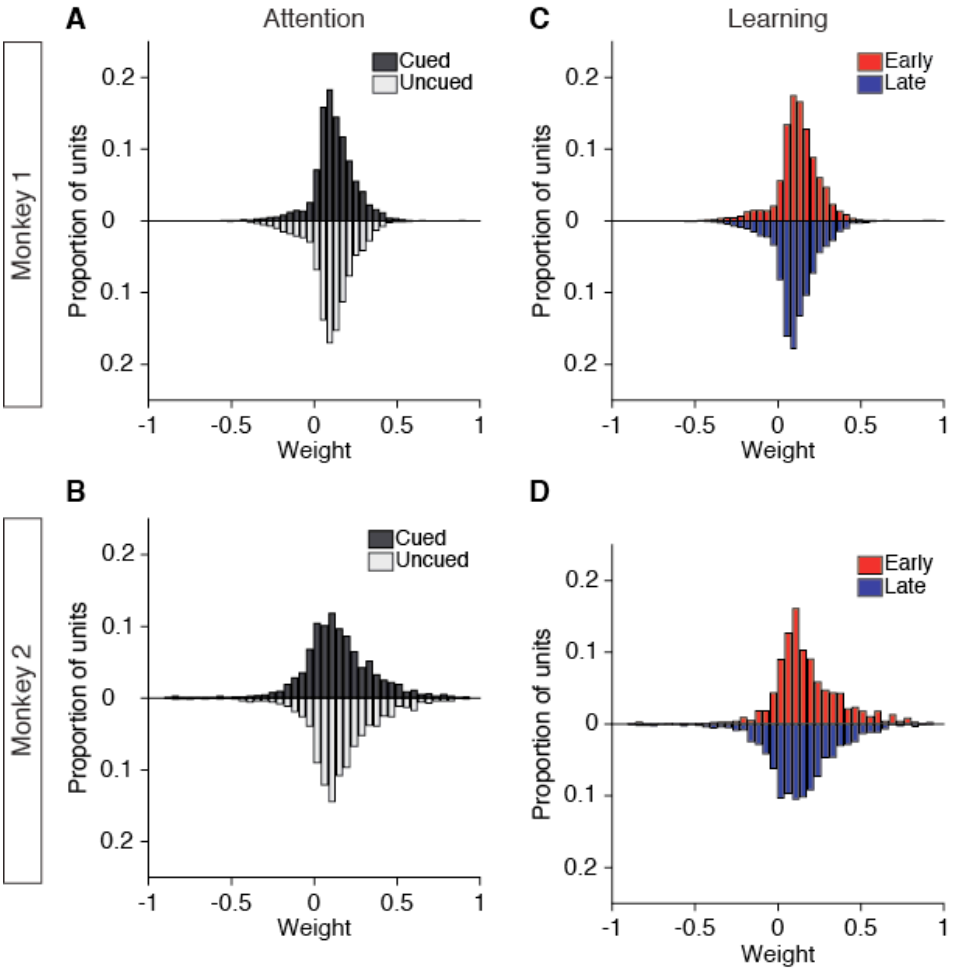


Fig. S8. The distribution of contributions of individual units to the first PC (correlated variability axis) is broad and not dominated by outlier units. Each histogram is the distribution of unit weights (the projection of each unit onto the 1st eigenvalue) for trials collected in the attention/learning condition in question. (**A**) Monkey 1 attention conditions: cued (top histogram) vs. uncued (bottom histogram). (**B**) Same for Monkey 2. (**C**) Monkey 1 perceptual learning conditions (sessions divided into first vs. second half of training): early learning (top histogram) vs. late learning (bottom histogram). (**D**) Same for Monkey 2. Additionally, the average firing rates of the units in the population were correlated with the projections of these firing rates onto the axis of the first PC, when analyzed across sessions per monkey (Monkey 1: Pearson correlation coefficient: $R = 0.25$; $p < 10^{-6}$; Monkey 2: $R = 0.37$, $p < 10^{-6}$).

Fig. S9

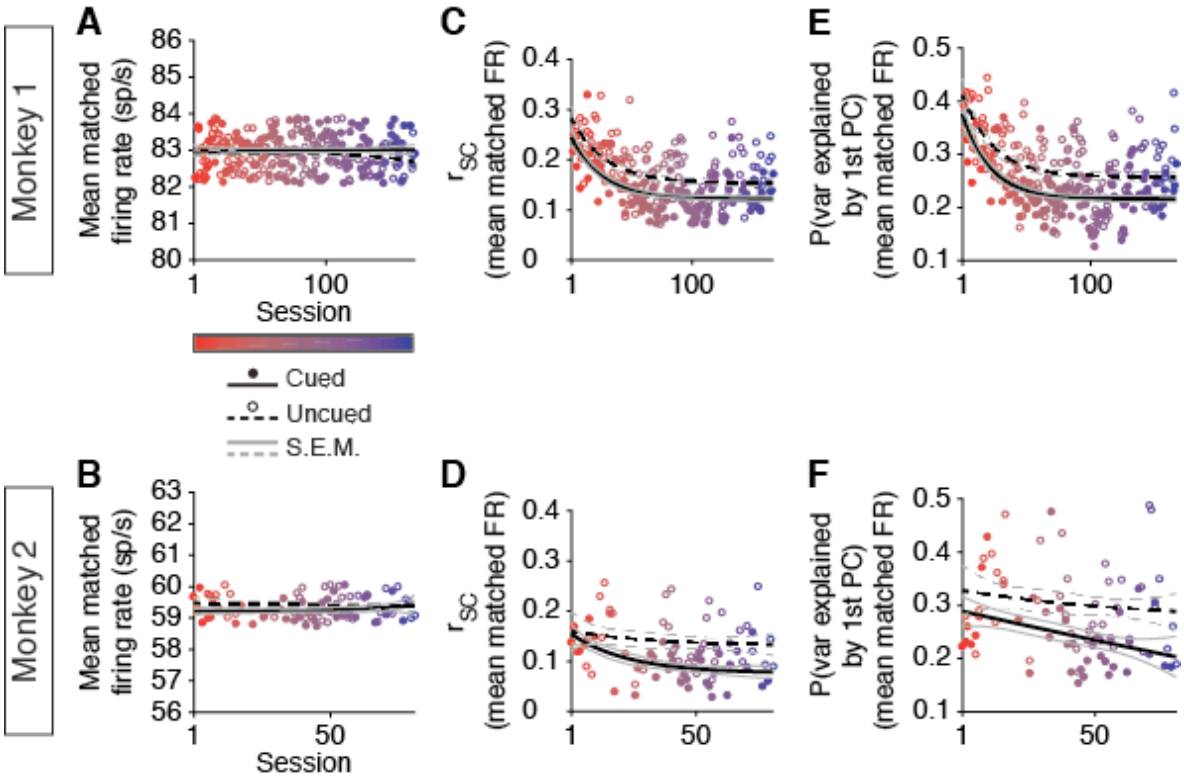


Fig. S9. The effects of attention and perceptual learning on correlated variability and the proportion of variance explained by the first principal component remain after firing rate mean matching. The firing rates of individual units are known to affect measurements of noise correlations (*2*). To control for such effects, we analyzed subsets of unit pairs such that the distributions of firing rates were matched across both attention conditions and all sessions. (**A**, **B**) The mean firing rates of units that survived this distribution matching procedure are, by definition, the same across sessions and attention conditions. (**C**, **D**) Matching firing rates did not affect our noise correlation results. The average r_{SC} decreased with both attention (Monkey 1: $p < 10^{-9}$; Monkey 2: $p < 10^{-3}$) and learning (Monkey 1: $p < 10^{-3}$; Monkey 2: $p < 10^{-3}$). (**E**, **F**) Matching firing rates did not affect the attention- and learning-dependence of the amount of variability explained by the first PC. The proportion of variance explained by the first PC decreased with both attention (Monkey 1: $p < 10^{-11}$; Monkey 2: $p < 10^{-3}$) and learning (Monkey 1: $p < 10^{-3}$; Monkey 2: $p < 10^{-3}$). Additionally, we analyzed subsets of units such that the distributions of Fano factor were matched across both attention conditions and both learning conditions (early vs. late learning). The average r_{SC} still decreased with both attention (Monkey 1: $p < 10^{-5}$; Monkey 2: $p < 10^{-5}$) and learning (Monkey 1: $p < 10^{-3}$; Monkey 2: $p < 10^{-3}$). The variability explained by the first PC still decreased with both attention (Monkey 1: $p < 10^{-6}$; Monkey 2: $p < 10^{-4}$) and learning (Monkey 1: $p < 10^{-3}$; Monkey 2: $p < 10^{-3}$).

Fig. S10

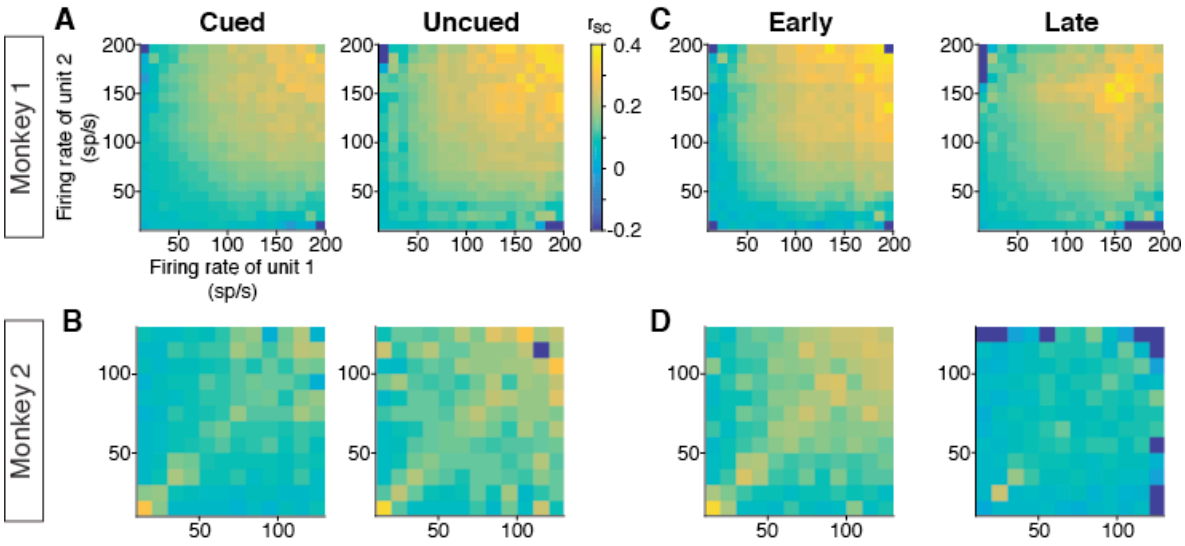


Fig. S10. Attention- and perceptual learning-related decreases in r_{SC} are not artifacts of changes in firing rate. Each plot is a heat map of r_{SC} as a function of the mean firing rates of the two units in each pair, and each plot is symmetric across the diagonal. (**A**, **B**) As expected, there is a relationship between firing rates and r_{SC} (*2, 3*). However, attention-related decreases in r_{SC} (uncued vs. cued trials) are broad and are not artifacts of changes in firing rate. (**C**, **D**) Similarly, learning-related decreases in r_{SC} (early vs. late trials) are broad and are not artifacts of changes in firing rate.

Fig. S11

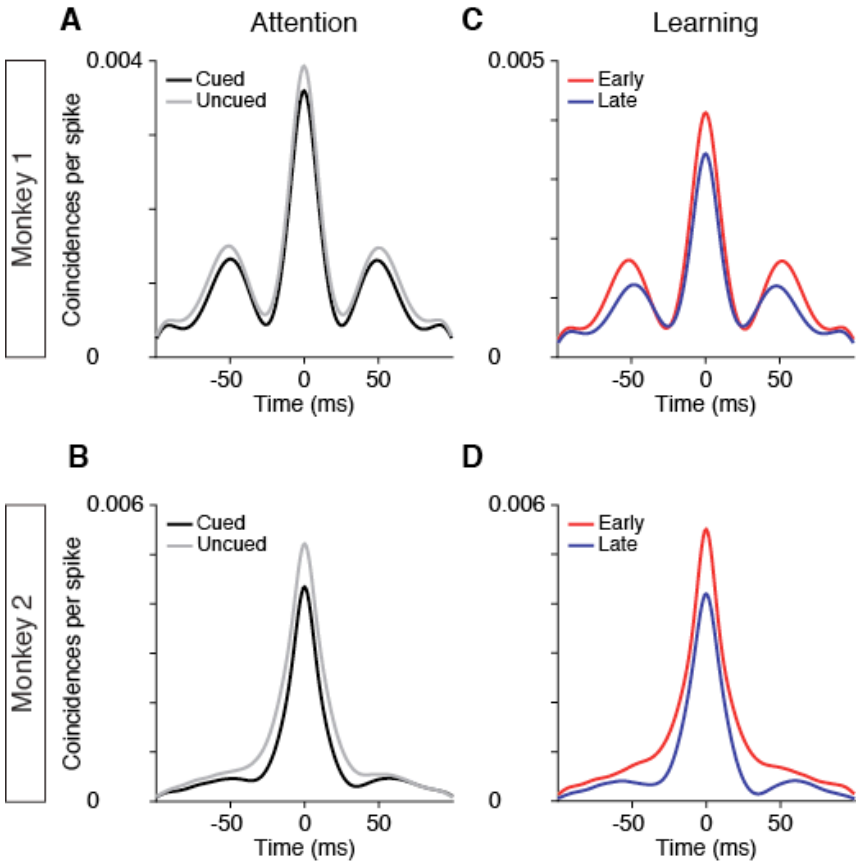
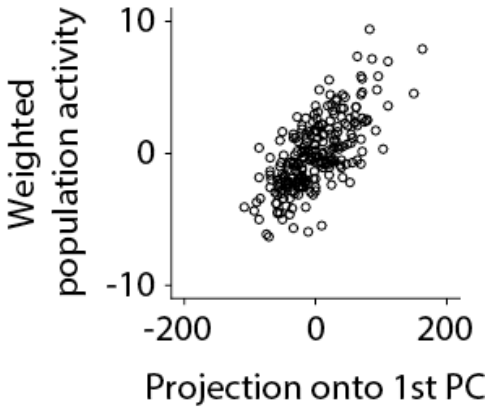


Fig. S11. Attention- and learning-related changes in correlated variability occurred across a broad range of timescales. The average shuffle corrected cross-correlogram (CCG) is plotted for each attention condition (**A**, **B**) or learning stage (**C**, **D**). The population average was smoothed with a 5 ms Gaussian kernel.

S12



S12. Comparison of the first principal component to weighted population activity. To investigate the relationship between the monkeys' choices and activity along the first PC in a complementary way to the choice decoder illustrated in fig. 4, A, E, and F, we compared projections of population responses to the stimuli before the orientation change onto the first PC with weighted sums of population activity using a method described by Haefner and colleagues (*22*) to infer the weights the monkeys used to make decisions. This plot illustrates the correlation between the projections onto the first PC and the weighted sums predicted by this decoding method for the data from one example day (Pearson correlation coefficient: $R = 0.70$, $p < 10^{-35}$). These two measures were highly correlated for both monkeys (Monkey 1: median Pearson correlation coefficient across all days: $R = 0.69$; two-tailed Wilcoxon signed rank test of the Pearson correlation coefficient across all days: $p < 10^{-8}$; Monkey 2: $R = 0.48$, $p < 10^{-6}$). These complementary methods show that there is a strong relationship between correlated variability and behavior on individual trials.

Fig. S13

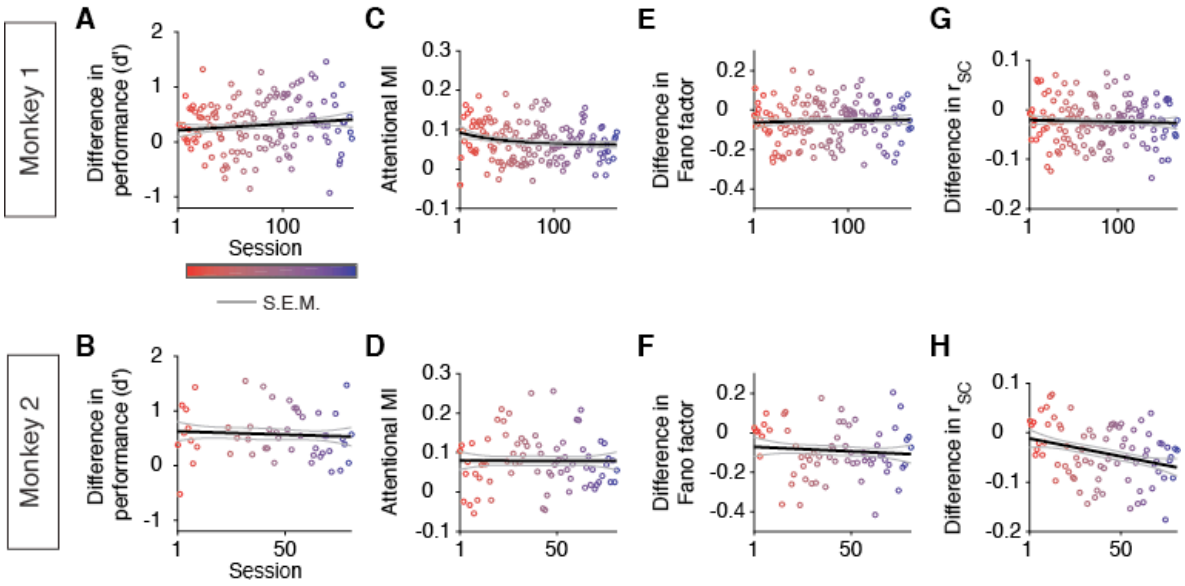


Fig. S13. The monkeys were not learning to attend during the recording period. In our study, the similarity in the neurophysiological effects of attention and perceptual learning is unlikely to be caused by the known interactions between attention and perceptual learning (*8*). Here we illustrate that perceptual learning was not accompanied by changes in the signatures of attention across sessions. Each detection task plot in fig. 2 illustrates both the within-session changes with attention (difference between the black solid and black dashed lines) and the across-session changes with learning (the exponential change in the black solid line across the x-axis). Here, we plot the difference between the black solid and black dashed lines from fig. 2 to illustrate whether the strength of attention effects changed across sessions. Changes across sessions quantified as per fig. 2. (**A**) The difference in behavioral sensitivity between attention conditions did not change across sessions for Monkey 1 ($p = 0.05$). (**B**) Same for Monkey 2 ($p = 0.55$). (**C**) The difference in evoked response between attention conditions, quantified with an attentional modulation index (MI) as previously described (*5*), did not change across sessions for Monkey 1 ($p = 0.38$). (**D**) Same for Monkey 2 ($p = 0.27$). (**E**) The difference in Fano factor between attention conditions did not change across sessions for Monkey 1 ($p = 0.50$). (**F**) Same for Monkey 2 ($p = 0.66$). (**G**) The difference in correlated variability (r_{SC}) between the attention conditions did not change across sessions for Monkey 1 ($p = 0.73$). (**H**) Same for Monkey 2 ($p = 0.30$). Number of sessions: Monkey 1: $n = 150$, Monkey 2: $n = 78$.

References and Notes

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

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

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

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

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

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

7. J. H. R. Maunsell, Neuronal mechanisms of visual attention. *Annu. Rev. Vis. Sci.* **1**, 373–391 (2015). doi:10.1146/annurev-vision-082114-035431 Medline

8. T. Watanabe, Y. Sasaki, Perceptual learning: Toward a comprehensive theory. *Annu. Rev. Psychol.* **66**, 197–221 (2015). doi:10.1146/annurev-psych-010814-015214 Medline

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

10. T. Kanashiro, G. K. Ocker, M. R. Cohen, B. Doiron, Attentional modulation of neuronal variability in circuit models of cortex. *eLife* **6**, e23978 (2017). doi:10.7554/eLife.23978 Medline

11. N. C. Rabinowitz, R. L. Goris, M. Cohen, E. P. Simoncelli, Attention stabilizes the shared gain of V4 populations. *eLife* **4**, e08998 (2015). doi:10.7554/eLife.08998 Medline

12. A. S. Ecker, G. H. Denfield, M. Bethge, A. S. Tolias, On the structure of neuronal population activity under fluctuations in attentional state. *J. Neurosci.* **36**, 1775–1789 (2016). doi:10.1523/JNEUROSCI.2044-15.2016 Medline

13. H. Nienborg, B. G. Cumming, Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* **459**, 89–92 (2009). doi:10.1038/nature07821 Medline

14. S. E. Kwon, H. Yang, G. Minamisawa, D. H. O'Connor, Sensory and decision-related activity propagate in a cortical feedback loop during touch perception. *Nat. Neurosci.* **19**, 1243–1249 (2016). doi:10.1038/nn.4356 Medline

15. J. M. Jeanne, T. O. Sharpee, T. Q. Gentner, Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron* **78**, 352–363 (2013). doi:10.1016/j.neuron.2013.02.023 Medline

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

17. J. F. Jehee, S. Ling, J. D. Swisher, R. S. van Bergen, F. Tong, Perceptual learning selectively refines orientation representations in early visual cortex. *J. Neurosci.* **32**, 16747–53a (2012). doi:10.1523/JNEUROSCI.6112-11.2012 Medline

18. D. A. Ruff, M. R. Cohen, Attention increases spike count correlations between visual cortical areas. *J. Neurosci.* **36**, 7523–7534 (2016). doi:10.1523/JNEUROSCI.0610-16.2016 Medline

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

20. C. Chandrasekaran, Computational principles and models of multisensory integration. *Curr. Opin. Neurobiol.* **43**, 25–34 (2017). doi:10.1016/j.conb.2016.11.002 Medline

21. T. E. Raghunathan, R. Rosenthal, D. B. Rubin, Comparing correlated but nonoverlapping correlations. *Psychol. Methods* **1**, 178–183 (1996). doi:10.1037/1082-989X.1.2.178

22. R. M. Haefner, S. Gerwinn, J. H. Macke, M. Bethge, Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat. Neurosci.* **16**, 235–242 (2013). doi:10.1038/nn.3309 Medline