Attentional fluctuations induce shared variability in macaque primary visual cortex

George H. Denfield,[1,*] Alexander S. Ecker,[1,2,3,4,*,†] Tori J. Shinn,[1] Matthias Bethge,[2,3,4] and Andreas S. Tolias[1,3,5]

[1] Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA
[2] Werner Reichardt Centre for Integrative Neuroscience and Institute of Theoretical Physics, University of Tübingen, Germany
[3] Bernstein Centre for Computational Neuroscience, Tübingen, Germany
[4] Max Planck Institute for Biological Cybernetics, Tübingen, Germany.
[5] Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA
[†] Lead Contact
* These authors have contributed equally

Corresponding author:

Alexander S. Ecker
Centre for Integrative Neuroscience
Otfried-Müller-Str. 25
72076 Tübingen, Germany
Email: alexander.ecker@uni-tuebingen.de
Phone: +49-7071-29889

Summary

Shared variability is common in neuronal populations, but its origin is unknown. Attention has been shown to reduce this variability, leading to the hypothesis that attention improves behavioral performance by suppressing common noise sources. However, even with precise control of the visual stimulus, the subject's attentional state varies across trials. While these state fluctuations are bound to induce some degree of correlated variability, it is currently unknown how strong their effect is, as previous studies have not manipulated the degree of attentional variability. Therefore, we designed a novel paradigm to dissociate changes in attentional strength from changes in attentional state variability and found a pronounced effect of attentional state fluctuations on correlated variability. This effect predominated in layers 2/3, as expected from a feedback signal such as attention. Thus, significant portions of shared neuronal variability may be attributable to fluctuations in internally generated signals, such as attention, rather than noise.

Keywords: spike count correlations, noise correlations, attention, primary visual cortex, V1, macaque, laminar probes

Introduction

Neuronal responses to repeated presentations of identical stimuli are highly variable (Softky and Koch, 1993). This variability can be correlated across populations of neurons (Bach and Krüger, 1986; Bair et al., 2001; Zohary et al., 1994), but its origin and significance is unclear.

One factor modulating correlations is attention. Studies of population activity in V4 found that attending to a stimulus inside the receptive fields of the recorded neurons reduced correlations in the trial-to-trial variability of the responses of those neurons to identical stimuli, compared to conditions in which attention was directed away from the receptive field (Cohen and Maunsell, 2009; Mitchell et al., 2009). These studies concluded that increasing the strength of attention reduces correlated variability by suppressing sources of shared noise (Fig. 1A).

However, because the subject's state of attention can be controlled on average but not precisely across trials, the strength of attentional modulation may vary from trial to trial even within a given attention condition (Cohen and Maunsell, 2010, 2011). Therefore, shared variability could also be driven by fluctuations in the state of attention (Fig. 1B). Indeed, the patterns of shared variability induced by fluctuations in the strength and spatial focus of gain-modulating signals such as attention are consistent with experimental data (Ecker et al., 2016; Rabinowitz et al., 2015).

In other words, correlated variability during attention tasks has been interpreted as evidence for both a suppression of common noise by attention (Cohen and Maunsell, 2009; Herrero et al., 2013; Mitchell et al., 2009) as well as trial-to-trial fluctuations of attentional state (Cohen and Maunsell, 2010, 2011; Ecker et al., 2016). However, it is unknown to what extent fluctuations in the state of attention indeed contribute to

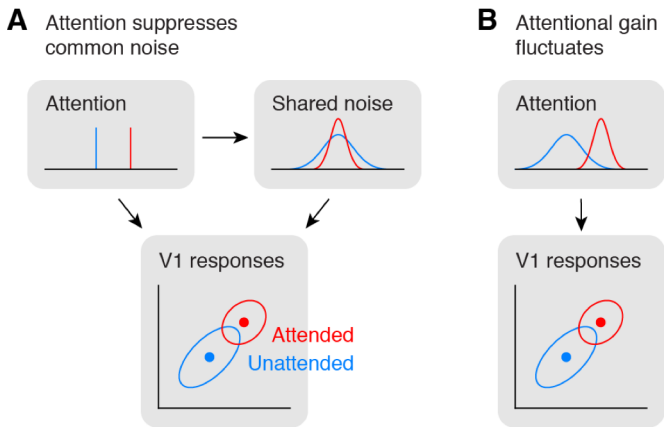

Figure 1. Attention and correlated variability. **A**) Hypothesis 1: Attentional gain is increased, but relatively stable under both conditions (top left). Correlated variability is driven by a common noise source (top right), which is suppressed by attention (Mitchell et al. 2009, Cohen & Maunsell 2009). **B)** Hypothesis 2: Attentional gain is increased, but fluctuates from trial to trial (Cohen & Maunsell 2010, 2011, Ecker at al. 2016). Correlated variability is driven by fluctuations of attentional state. The reduction in correlations under attention would imply that the attentional gain is less variable when attending.

correlated variability in population responses, because the paradigms employed in previous studies did not manipulate the degree of attentional fluctuations behaviorally.

To create such a scenario, we developed a novel, cued change-detection task that can dissociate changes in the strength of attention from changes in the variability of the attentional state by manipulating the behavioral relevance of two simultaneously displayed stimuli across task conditions. If the dominant factor governing levels of correlated variability is attentional suppression of common noise, we expect correlations to decrease as attentional strength increases, resulting in intermediate levels of correlations when both stimuli need to be attended (Fig. 2A). Alternatively, if fluctuations in attention are the dominant factor modulating correlations, correlations should be highest when both stimuli need to be attended (Fig. 2B), as this is the condition where attentional fluctuations are most likely to occur. In practice, of course, both mechanisms may contribute. However, the degree to which attentional fluctuations are relevant is revealed by considering the difference in correlations between conditions in which attention is focused on a single stimulus and when attention to both stimuli is required (Fig. 2C).

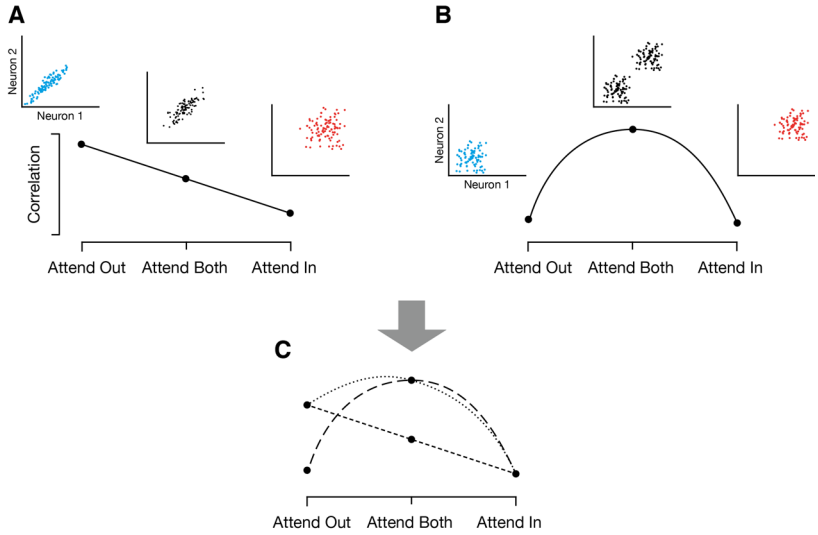

Figure 2. Predicted effects of attention on correlations when attending one ("Attend In/Out") or two stimuli ("Attend Both"). **A)** Scenario in which attentional fluctuations are negligible and attention primarily acts by suppressing common noise sources. In this case, we expect intermediate correlations when attending two stimuli. **B,)** Scenario in which fluctuations in attention induce correlations. In this case, we expect attention to switch randomly between the two targets in the "Attend Both" condition, resulting in the highest correlations in this condition. **C)** Both mechanisms may contribute to different degrees. The relevance of attentional fluctuations is revealed by the relative level of correlation in the "Attend Both" condition.

We recorded neuronal responses from primary visual cortex of macaque monkeys while they performed this task and found that attention modulated firing rates of V1 neurons. We found that shared variability was highest when both stimuli were behaviorally relevant and lowest in conditions in which only one stimulus was the focus of attention, arguing that fluctuations in the state of attention, induced by changes in attentional allocation strategies, are an important factor governing shared neuronal variability. This modulation predominated in supragranular cortical layers, as expected if it were due to a feedback signal such as attention (Anderson and Martin, 2009; Maunsell and van Essen, 1983; Rockland and Pandya, 1979; Ungerleider et al., 2008).

Results

Change detection task and manipulation of attentional allocation strategy

We trained two rhesus macaque monkeys to perform a cued, orientation-change detection task (Fig. 3A). A trial was initiated when the subject fixated a central fixation spot. Two "noisy" Gabor patches appeared symmetrically in the lower left and lower right visual field 300ms later. During the Zero-Coherence Period (ZCP), these patches randomly changed their orientation every frame (10ms per frame; set of 36 orientations evenly spaced between 0 and 175 degrees). After a random period of time, drawn from an exponential distribution (minimum of 0.01s, mean of 2.17s, and maximum of 5s), one of the two stimuli entered the Coherent Period (CP). During the CP one particular orientation, called the "signal" orientation, was shown with a higher probability than the other orientations. By varying this probability, we could control the "coherence" of the stimulus, making the occurrence of the signal orientation more or less obvious over the background orientation noise, to manipulate the difficulty of a trial. The occurrence of this signal orientation was the change the monkey had to detect, which he reported by making a saccade to the changed stimulus within a short reaction time window. On 10% of trials no signal orientation occurred, and the monkey was rewarded for maintaining fixation throughout the trial.

We used a cued block design to manipulate the focus of the subject's attentional state (Fig. 3B), where the cue was the color of the fixation spot. Two of these conditions, "attend in" (AI) and "attend out" (AO), were similar to those in typical spatial attention tasks, where the stimulus overlapping the neurons' receptive fields is cued in the AI condition, and the other

stimulus is cued in the AO condition. The cues for these conditions (red for AI, blue for AO) were 100% valid, such that the change occurred only at the cued location. In the condition labeled "attend both" (AB), indicated by a black fixation spot, either stimulus had an equal probability (50%) of showing the change on a given trial.

Our paradigm therefore differs from typical covert attention tasks used to study neuronal variability in two respects. First, during the AI and AO conditions in our task, there are no catch trials with invalid cues (Cohen and Maunsell, 2009) or signals in the distractor that need to be ignored (Herrero et al., 2013). While catch trials are typically used to measure the behavioral shift due to attention, they are likely to induce attentional fluctuations, as they render the cue unreliable and encourage some degree of attentional focus on the non-cued stimulus by rewarding successful performance at that location. As our goal in the AI and AO conditions is to minimize attentional fluctuations, we used 100% reliable cues. In our third condition (AB),

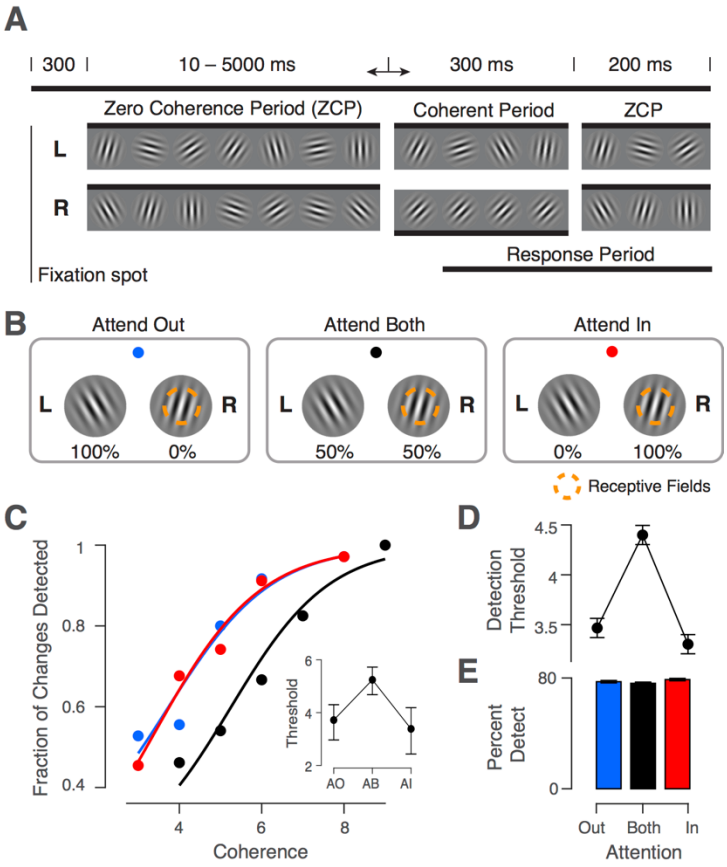

Figure 3. Task diagram with behavioral results. **A**) Orientation change-detection task. Two stimuli (L: left, R: right) randomly change their orientation during the ZCP (length 10-5000ms). One stimulus (R in this example) then enters the CP (300ms) when the signal orientation is shown (coherence exaggerated for clarity). This period is followed by another 200ms ZCP to allow time for a behavioral response. **B**) Illustration of attention conditions. Attention is cued according to fixation spot color. This color scheme is used in all figures to represent each condition. Percentages below the stimuli indicate the probability that the change occurs in this stimulus on a given trial. One stimulus overlaps the recorded neurons' receptive fields. **C**) Example session psychophysical performance. Individual points represent fraction of changes detected at a given coherence. Solid lines indicate fit of logistic function to the data. Inset shows detection threshold with 95% CIs. **D**) Behavioral summary. Same as inset in **c**, but averaged across sessions in our dataset (N=27; mean±SEM). **E)** Percentage of changes detected in each condition averaged across sessions (mean±SEM).

indicated by a black fixation spot, either stimulus was equally likely to change. We used this condition as the baseline to measure the behavioral improvement attributable to attention, analogous to how other paradigms use catch trials.

There were, therefore, three attentional conditions but two attentional strategies that our task engaged. To maximize reward in the AI and AO conditions, attention should be focused on only the cued stimulus. With attention deployed consistently across trials with regard to spatial location, attentional state fluctuations should be minimized. In the AB condition, attention should fluctuate more strongly between the two spatial locations across trials, as ignoring one of the stimuli is no longer a viable strategy for maximizing reward. One way to conceive of this allocation strategy is that the AB condition is comprised of a mixture of the attentional states deployed in the AI and AO conditions. Note, attentional state fluctuations need not be non-existent in the AI and AO conditions but only decreased relative to the AB condition in order to test our hypothesis.

If subjects used the strategies described above, there should be some trials in the AB condition where the subject attended the unchanged stimulus and required a higher coherence level to notice a change in the correct stimulus on that trial. Such occurrences would lead to a rightward shift in the psychometric function and higher detection thresholds in the AB condition. The example session in Figure 3C exhibits a clear rightward shift in the psychometric curve along with a significantly elevated coherence threshold in the AB condition. This effect was consistent across sessions (Fig. 3D, $p < 10^{-10}$, two-way ANOVA), being present in 22 out of 27 sessions. To avoid potential confounds from changes in task difficulty across attention conditions, we balanced the overall percent correct performance in each condition by raising coherence levels one step in the AB condition. Overall, subjects identified an average of 77±1.2% of changes, and there was no significant effect of attention condition on performance (Fig. 3E, $p = 0.10$, two-way ANOVA).

Our goal was to develop a behavioral paradigm in which attention could fluctuate or shift between two stimulus locations – the AB condition – and remain focused on one location in the other conditions. While our behavioral results are consistent with this attentional allocation strategy, they are also consistent with a strategy in which attention acts as a zoom lens, as

suggested in Eriksen and St James (1986), widening its focus to encompass both stimuli simultaneously. Note, the fact that detection thresholds are elevated in the AB condition suggests that if attention is allocated to both stimuli simultaneously, the stimuli are not processed to the same degree as they are in the AI or AO conditions. That is, widening the attentional field entails a reduction in attentional strength within the field. As we will see, however, these strategies make different predictions for the patterns of correlated variability we expect to see across our task conditions.

Attentional modulation of neuronal firing rates

While subjects performed the task, we recorded spiking responses from neurons in primary visual cortex using 32-channel silicon probes with a spacing of 60μm between channels (NeuroNexus V1x32-Edge-10mm-60-177). We recorded 416 single units (15.4±.95 units per session) across 27 sessions (N=7 from Subject B, N=20 from Subject D) from two male macaque monkeys. The two Gabor stimuli in our task were placed symmetrically in the lower visual field with one stimulus covering the receptive fields of the recorded neuronal population. Given the laminar nature of our recordings, receptive fields overlapped almost completely.

Our highly dynamic stimulus drove neurons strongly, with mean firing rates of 23.2±1.1 spikes/sec across sessions. Consistent with previous studies we found that attention increased firing rates of V1 neurons (McAdams and Maunsell, 1999; McAdams and Reid, 2005; Motter, 1993; Roelfsema et al., 1998), with on average ~30% of single units being significantly modulated by attention in a given session. This modulation was present in both the AI and AB conditions and appeared strongest early in the ZCP (Fig. 4A).

Note, our dataset contains fewer trials of long duration, given the exponential distribution of ZCP lengths and a slight tendency of subjects to prematurely abort longer trials (only ~40% of valid trials are longer than 1s, and ~15% are longer than 2s). We thus focused our analyses on the first second after stimulus onset, in which attentional modulation of firing rates was strongest, and on correct trials, where we can have the most confidence that the subject's attention was oriented as desired in our task. Additionally, note that all analyses of firing rates and spike counts were performed during the ZCP, before any changes in stimulus coherence or behavioral

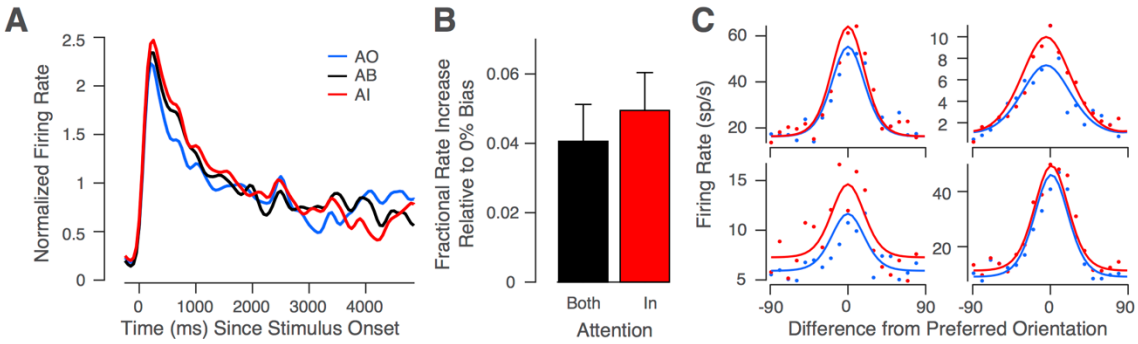

Figure 4. Attentional modulation of neuronal responses **A)** Example session spike density function for each condition, normalized to the average response in AI condition (mean across units). **B)** Fractional increase in firing rates in the AB and AI conditions relative to the AO condition averaged across sessions (N=27; mean±SEM). **C)** Example single unit tuning curves in AI (red) and AO (blue) conditions. Dots show responses to specific orientations; solid lines show fitted von Mises functions.

responses were made, ensuring that analyses were performed on identical stimuli across conditions.

We first calculated fractional firing rate increases in the AI and AB conditions, relative to the AO condition (Fig. 4B). During this interval, firing rates in the AI and AB conditions were significantly elevated relative to the AO condition (AI: 5.0±1.1% increase, $p = 0.0001$, Bonferroni-corrected t-test, $\alpha=0.0167$; AB: 4.1±1.1%, $p = 0.001$) but not different from each other ($p = 0.35$). Amongst the roughly 30% of units showing significant modulation of firing rates by attention, around 32% showed pure gain modulation, around 20% showed pure offset modulation, while the remainder exhibited a mixture of multiplicative and additive modulation. Examples of pure gain- versus pure offset-modulated cells are shown in Figure 4C.

Differentiating the effects of attention on shared variability

Our behavioral and neurophysiological results so far, beyond demonstrating that our task engages attention, are consistent with two different attentional allocation strategies in the AB condition, while we conclude that attention is primarily focused on the single, relevant stimulus in the AI and AO conditions. The first strategy involves widening the focus of attention to encompass both stimuli. In this case, we would expect attentional fluctuations to be negligible. This scenario would support the interpretation that attention suppresses a common noise source (Mitchell et al. 2009, Cohen & Maunsell 2009) and we would expect correlations to be intermediate in the AB condition (Fig. 2A). The second strategy involves shifting the focus of attention

randomly between the two stimuli. In this case, we would expect correlations to be highest in the AB condition (Fig. 2B). Note that this scenario does not rule out the possibility that attention suppresses a common noise source, as both mechanisms could be at play (Fig. 2C). However, given that the same dataset has been interpreted as evidence that attention suppresses noise (Cohen & Maunsell 2009) and that attention fluctuates (Cohen & Maunsell 2010), it is an important question to quantify to what degree attentional fluctuations induce trial-to-trial variability.

Attentional modulation of shared variability

To measure the degree to which attentional fluctuations induce trial-to-trial variability, we calculated pairwise spike count correlations over repeated presentations of identical ZCP sequences in each attention condition. Our results match the predictions in Figure 2B and support the hypothesis that fluctuations in the state of attention are the dominant factor inducing shared neuronal response variability in our dataset (Fig. 5A). Spike count correlations were significantly modulated by attention condition ($p = 0.00002$, two-way ANOVA), correlations were highest in the AB condition ($p = 0.00001$, t-test, see methods), and correlations in the AI and AO conditions were not significantly different from one another ($p = 0.82$, post-hoc Tukey's test). This relationship held individually for both subjects (Fig. 5B "task"; Subject B: $p = 0.013$, Subject D: $p = 0.002$, two-way ANOVA). Task-evoked correlations were higher overall in Subject D than in Subject B, though both subjects had more comparable correlation levels during fixation when no stimulus was present (Fig. 5B "fix"). Despite a clear modulation of shared variability across attention conditions, Fano factors, a measure of individual neuronal variability, assessed over the same time interval were not modulated significantly by attention condition ($p = 0.21$, two-way ANOVA). However, this result is likely due to a lack of statistical power, because the estimation error for Fano factors was larger than the expected effect given the correlation differences.

Additionally, fixational eye movements, also called micro-saccades, cannot account for our results, as there was no difference in the number of such events across attention conditions ($p = 0.25$, two-way ANOVA). Note also that these results are not trivially explained by changes in firing rates across conditions, as firing rates in the AI condition were elevated compared to the

AO condition (Fig. 4B), but correlation magnitudes were not significantly different in these conditions (Fig. 5A and B). Nor do changes in stimulus coherence function as an explanation for elevated correlations in the AB condition, as spike counts were analyzed during the ZCP before any changes in the coherence of the stimulus occurred.

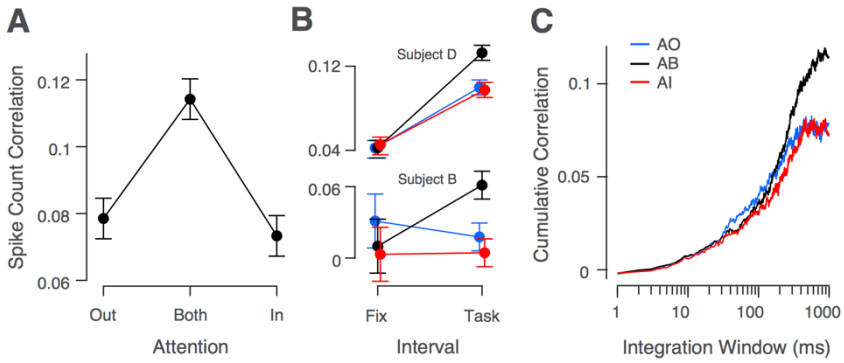

Figure 5. Effects of attention on shared variability. **A**) Spike count correlations from 0-1s following stimulus onset, averaged across sessions (N=27). **B**) Spike count correlations shown separately for both subjects during fixation (300ms interval) and during the task (same interval as in A). **C**) Cumulative correlation coefficient, calculated by integrating the cross-correlogram, for each attention condition and averaged across sessions. Data in A-B show mean ± SEM, C omits SEM.

Next, we wanted to investigate the timescale of the correlation effect we found, to better understand its origin. Synaptic processes unfold on the millisecond scale whereas cognitive processes, such as attention, unfold over longer timescales. Behavioral work suggests that voluntarily shifting attention between different stimuli takes on the order of several hundred milliseconds (Duncan et al., 1994; Müller et al., 1998). Thus, if attention is indeed shifting between the two stimulus locations during the AB condition, these psychophysical results provide a lower bound for the timescale over which we expect to see correlations rise in the AB condition.

Using the relationship between spike count correlations and cross-correlograms described in Bair et al. (2001) and modified in Ecker et al. (2014), we calculated spike train cross-correlograms for neuronal pairs in each attention condition and integrated them from 1ms to 1000ms, our maximum counting window. Examining the point at which the resulting correlation levels saturate provides an estimate of the timescale of correlation. The results in Figure 5C show that correlations in the AB condition began to diverge from the AI and AO conditions after 200ms, and correlations in the AI and AO condition saturated to similar levels near 400ms, while AB correlations continued to rise for several hundred milliseconds more. The time course of these results fits well with the estimated time course of changes in attentional state (Duncan et al., 1994; Müller et al., 1998). Interestingly, between 40ms and 400ms, the level of correlations appeared

lower in the attended versus unattended conditions (Fig. 5C), consistent with earlier work (Cohen and Maunsell, 2009; Herrero et al., 2013; Mitchell et al., 2009) and suggesting that attention indeed suppresses common noise at this faster timescale. However, despite being consistent with previous results and being observable in both monkeys individually (data not shown), this trend was not statistically significant in our dataset ($p = 0.074$ at 100ms, two-way ANOVA without correction for multiple comparisons).

Laminar profile of attention effects

To examine the laminar profile of the attentional modulation of firing rates and shared variability, we calculated the current source density (CSD; Mitzdorf, 1985) across channels for each session from the task-stimulus evoked local field potentials (Fig. 6A). These profiles were quite consistent across sessions, with the most prominent stimulus-evoked sink-source configurations in L5-6 and L1-2/3, largely washing out the earliest sink-source switch typical of the L4-5 boundary (van Kerkoerle et al. (2017) report a similar effect). We computed CSDs to aid in the grouping of single units into the supragranular (S), granular (G), or infragranular (I) layers, but we also took advantage of known electrophysiological characteristics of cells in different layers (Snodderly and Gur, 1995; see methods). The most reliable such property was the high spontaneous activity associated with L4C (Snodderly and Gur, 1995), which was readily discernible from multi-unit activity and was located consistently close to the L4-5 boundary determined from the CSD. Additional factors included the weaker orientation tuning of the deep granular layer and smaller receptive fields (Fig. 6A). The first channel below the L4-5 boundary was our zero-point for relative unit depths. We defined the granular layer as the first 400μm superficial to the L4-5 boundary, consistent with previous histological (Fitzpatrick et al., 1985; Lund, 1988) and recent electrophysiological studies (Hansen et al., 2012; Smith et al., 2013). All units above this 400μm band were labelled supragranular, and all those below it were labelled infragranular. The G-I (L4-5) boundary could be determined most reliably across sessions, but the S-G boundary could not always be determined as precisely. We therefore varied the cut-off boundary between the supragranular and granular groups over a span of nearly 200μm and re-calculated the results presented in Figure 6. Doing so did not qualitatively affect our results.

Attentional modulation of V1 neuronal responses is thought to be a feedback process (Buffalo et al., 2009; Buschman and Miller, 2007; Gregoriou et al., 2009), and anatomical work has shown that feedback projections from higher order visual areas target the supra- and infra-granular layers (Anderson and Martin, 2009; Maunsell and van Essen, 1983; Rockland and Pandya, 1979; Ungerleider et al., 2008). As a result, we expected the strongest attentional modulation of firing rates to manifest there. Indeed, in the supragranular group, firing rates were modulated most strongly (Fig. 6B), and this modulation was significant in both the AB and AI conditions relative to the AO condition (AB: 5.5±1.2%, p = 0.0001, AI: 5.5±1.3%, p = 0.0004, Bonferroni-corrected t-test, α=0.025). In the infragranular group, there was also significant modulation of firing rates in the AI condition but not the AB condition (AB: 3.4±1.6%, p = 0.045, AI: 5.5±2.0%, p = 0.011, α=0.025). In the granular group, firing rates were not significantly elevated in the AB or AI conditions (.45±1.7%, p = 0.814, AI: 4.0±1.7%, p = 0.035, α=0.025).

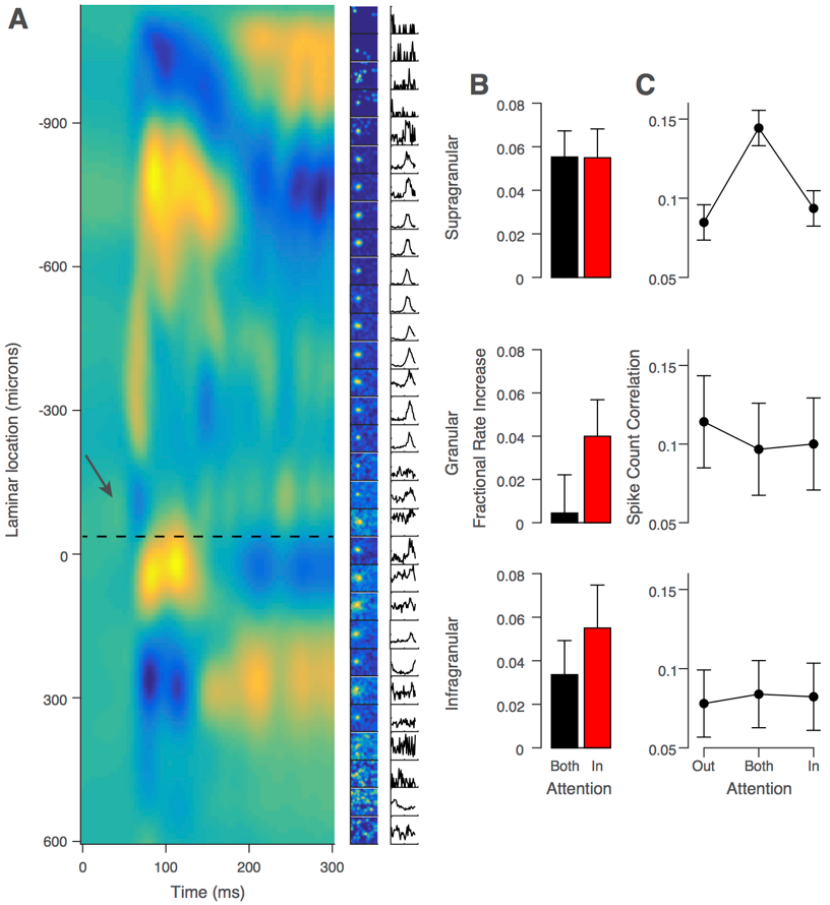

Figure 6. Laminar profile of attention effects. **A**) Example session CSD profile evoked by task stimulus (left column) with multi-unit receptive fields (middle) and tuning curves (right). Depths are relative to first L5 channel. Dotted black line shows L4-5 transition. Arrow shows initial current sink-source flip in L4C. **B**) Fractional increase in firing rates in AB and AI, relative to AO, conditions split by laminar group. **C**) Spike count correlation over 0-1s interval split by laminar group. Data in B-C show mean across sessions ± SEM (N=27).

Next, we examined the laminar profile of attentional effects on spike count correlations (Fig. 6C). Correlations were significantly modulated by attention condition in the supragranular

group ($p = 0.0007$, two-way ANOVA). Post-hoc testing again showed correlations were highest in the AB condition ($p = 10^{-6}$) and equivalently low in the AI and AO conditions ($p = 0.84$). In the granular and infragranular groups, correlations were constant across attention conditions. Although there was a downward trend in overall spike count correlation magnitude from superficial to deep, there was no significant effect of layer ($p = 0.62$, two-way ANOVA; S: $r_{sc} = 0.11 \pm 0.02$, G: $r_{sc} = 0.10 \pm 0.03$, I: $r_{sc} = 0.08 \pm 0.02$).

Discussion

We developed a task to dissociate changes in the strength of attentional modulation from changes in variability in the attentional state by varying the behavioral relevance of two simultaneously presented stimuli and encouraging the use of different attentional allocation strategies across task conditions. We found levels of shared variability to be highest in the condition in which both stimuli were behaviorally relevant, supporting the idea that this condition introduced competition for attentional resources, which increased attentional state variability. In contrast, shared variability was lowest in the conditions in which attention could be focused on only one stimulus. These results support the hypothesis that fluctuations in the state of attention can be a prominent source of shared neuronal response variability. More specifically, our results for correlations on timescales on the order of individual trials are most consistent with the scenario presented in Figure 2B, in line with the predictions of Ecker et al. (2016). More generally, our results suggest that a significant fraction of shared variability in neuronal populations can be attributed to fluctuations in behaviorally-relevant, internally generated signals, rather than shared sensory noise (Ecker and Tolias, 2014; Ecker et al., 2010, 2014, 2016; Goris et al., 2014; Haefner et al., 2016; Nienborg and Cumming, 2009; Rabinowitz et al., 2015).

We focused primarily on the level of correlations in the AB condition, the condition in which the two possible mechanisms driving correlations made diverging predictions (Fig. 2). Previous studies have focused on the comparison of AI vs. AO conditions and found a reduction of correlations when attending (AI condition). At the one-second timescale we analyzed, we did not observe this reduction, suggesting that the difference between AI and AO conditions observed in previous studies was not driven by fluctuations in attention. However, at a timescale around

100ms, we did find a trend towards lower correlations, consistent with earlier work that considered faster timescales (Cohen and Maunsell, 2009; Herrero et al., 2013; Mitchell et al., 2009). Taken together, these results suggest that both mechanisms – suppression of common noise and attentional fluctuations – are at play, but operate at different timescales. Also consistent with this picture, Verhoef and Maunsell (2017) recently proposed that the reduction of correlations under attention is due to a suppression of (variable) normalizing inputs from the unattended surround, which predicts that this effect should be limited to timescales of synaptic integration (i.e. < 100ms).

Because the impact of variability in the attentional state on correlations manifested on a timescale of individual trials in our task, should we therefore expect that fluctuations in internal signals, in general, only induce correlations on long timescales? For correlations resulting from fluctuations in a gain-modulating signal, correlations are roughly proportional to the square of the number of spikes in the count window (Ecker et al., 2010, 2016; Goris et al., 2014), so it is not until this window grows sufficiently large that the AB condition effect manifests. Ultimately, however, this timescale is likely to depend on the mechanism by which such signals impact neuronal populations. Work on orienting of attention and attentional dwell time suggests that voluntarily shifting attention between different stimuli takes on the order of several hundred milliseconds (Duncan et al., 1994; Müller et al., 1998). In our case, this shifting of attention between stimulus locations is the strategy we were hoping to induce in our paradigm and appears to be the likeliest explanation for how attention is allocated across trials in our AB condition, given our behavioral and neurophysiological results. We would, thus, expect that AB correlations should be elevated on a timescale of at least several hundred milliseconds, which is what we found (Fig. 5C).

Note that this line of reasoning stands regardless of whether the shift in attention that occurs involves a narrowly-focused attention field encompassing only one stimulus at a time – resembling the spotlight or narrowly-focused zoom lens models (Eriksen and St James, 1986; Eriksen and Yeh, 1985) – or whether some degree of attention is allocated to both stimuli simultaneously, but with one stimulus receiving a greater degree of attention than the other on a given trial – resembling the Variable Precision model of resource allocation (van den Berg et al., 2012). In this latter case, the shift of attention corresponds to alternations in which stimulus

receives the greater strength of attentional focus on a given trial. The key, however, is that some change in attentional resources allocated to the receptive field stimulus occurs across trials. Therefore, our results are not consistent with models of attention that suggest that both stimuli are processed simultaneously and that a consistent or uniform degree of attentional processing is distributed across the full field of attention.

Recent studies have examined the laminar profile of attentional modulation of firing rates (van Kerkoerle et al., 2017) or of spike count correlations during passive fixation (Hansen et al., 2012; Smith et al., 2013). Only one study has examined the laminar relationship between attentional modulation and shared variability (Nandy et al., 2017), and ours is the first to do so in V1. Nandy et al. (2017) found significant attentional modulation of firing rates in all layers, with the strongest effects in the granular layer. In contrast, van Kerkoerle et al. (2017) found the weakest attentional modulation of firing rates in the granular layer of V1. Our results are in better agreement with those of van Kerkoerle et al. (2017), as we found the strongest attentional modulation of firing rates in the supragranular, followed by the infragranular layers, as expected given the anatomical distribution of feedback cortical connections (Anderson and Martin, 2009; Maunsell and van Essen, 1983; Rockland and Pandya, 1979; Ungerleider et al., 2008).

Regarding correlation magnitude across layers, we observed a different pattern of results from both Nandy et al. (2017), who found the highest correlations in the granular layer of V4, and Hansen et al. (2012) and Smith et al. (2013), who found the lowest correlations in the granular layer in V1. In our study, overall correlation magnitude did not differ significantly by layer. These differences across studies could be attributable to the variable behavioral demands placed on each study's subjects, which would be consistent with our overarching hypothesis that differences in correlation magnitude across studies can be accounted for in large part due to differences in the nature of the internal signals engaged by different tasks and how they are deployed to meet the subjects' behavioral needs.

Indeed, we created a task in which one condition's behavioral demands were quite different (AB) from those of the other two conditions (AI, AO), and we found a large difference in correlations that varied with those demands, which was confined primarily to the supragranular layers. This modulation of correlations was not present in the infragranular layers, despite

attentional modulation of rates in the AI condition. One reason may be a lack of sufficient statistical power. Most of our isolated single units were from the supragranular layers (just over eight units per session on average), with about half that number isolated in the infragranular layers, and fewer still from the granular layer. The difference could also be attributable to the anatomical and computational characteristics of each layer, which by no means are completely understood (Callaway, 1998; Douglas and Martin, 2004; Lund, 1988). The infragranular layers additionally receive feedback from and send projections to subcortical regions (Lund et al., 1975) and such signals may modulate shared variability differently. Ultimately, the finding that fluctuations in attention predominantly modulate correlations in the supragranular layers matches the location where we found the most pronounced attentional modulation of firing rates and accords well with the known anatomy of corticocortical interactions, particularly for feedback signals.

Finally, there has been an increasing interest in recent years in leveraging population recording and latent-variable modeling techniques to infer the state of internally-generated, cognitive signals, such as attention, on more behaviorally-relevant timescales, to better understand the nature of these signals and their impact on decision-making and behavior (Afshar et al., 2011; Engel et al., 2016; Latimer et al., 2015; Rabinowitz et al., 2015; Yu et al., 2009). To make such inferences, these methods make use of the patterns of covariance in population activity and rely on the assumption that this variability occurs in a low-dimensional space (e.g., the "attention axis" of Cohen and Maunsell (2010)). A further, but critical, assumption of these techniques is that much of this shared variability is not noise but is attributable to the action of behaviorally-relevant, internally generated signals. However, a clearer demonstration that changes in internal signals indeed contribute significantly to shared neuronal variability was lacking. We presented a paradigm designed specifically to test for such a contribution, and our results provide support for this critical assumption. Additionally, our results demonstrate the subtlety of the effects that internal signals such as attention have on correlated variability, exemplified by the two timescales over which attention modulated correlations.

Author Contributions

Conceptualization, G.H.D, A.S.E and A.S.T; Methodology, G.H.D, T.J.S, A.S.E, A.S.T; Software and Validation, G.H.D and A.S.E; Formal Analysis, G.H.D; Investigation, G.H.D, T.J.S; Resources, A.S.E, M.B and A.S.T; Writing – Original Draft, G.H.D; Writing – Review & Editing, G.H.D, A.S.E, T.J.S, M.B and A.S.T; Visualization, G.H.D; Supervision, A.S.E, M.B and A.S.T; Funding Acquisition, A.S.E, M.B and A.S.T.

Acknowledgments

We thank Amy M. Morgan and Camila Lopez for technical assistance and Dimitri Yatsenko for discussion and the development of DataJoint. This work was supported by grants NEI R01-EY018847-05, NEI R01-EY026927-01A1, NEI P30-EY002520-33 and the NIH-Pioneer award DP1-OD008301 to A.S.T. This work was also supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior/Interior Business Center (DoI/IBC) contract number D16PC00003. The US Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/IBC or the US Government; German Research Foundation (DFG) grant EC 479/1-1 to A.S.E; the Bernstein Center for Computational Neuroscience (FKZ 01GQ1002); the German Excellency Initiative through the Centre for Integrative Neuroscience Tübingen (EXC307); G.H.D was supported by NEI T32-EY007001-40, Baylor College of Medicine (BCM) and the BCM Medical Scientist Training Program. The authors have no conflicts of interest to report.

References

Afshar, A., Santhanam, G., Yu, B.M., Ryu, S.I., Sahani, M., and Shenoy, K.V. (2011). Single-trial neural correlates of arm movement preparation. Neuron *71*, 555–564.

Anderson, J.C., and Martin, K.A.C. (2009). The Synaptic Connections between Cortical Areas V1 and V2 in Macaque Monkey. Journal of Neuroscience *29*, 11283–11293.

Bach, M., and Krüger, J. (1986). Correlated neuronal variability in monkey visual cortex revealed by a multi-microelectrode. Experimental Brain Research *61*, 451–456.

Bair, W., and O'Keefe, L.P. (1998). The influence of fixational eye movements on the response of neurons in area MT of the macaque. Vis. Neurosci. *15*, 779–786.

Bair, W., Zohary, E., and Newsome, W.T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. The Journal of Neuroscience *21*, 1676–1697.

van den Berg, R., Shin, H., Chou, W.-C., George, R., and Ma, W.J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. Proceedings of the National Academy of Sciences *109*, 8780–8785.

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

Buffalo, E.A., Fries, P., Landman, R., Liang, H., and Desimone, R. (2009). A backward progression of attentional effects in the ventral stream. Proceedings of the National Academy of Sciences *107*, 361–365.

Buschman, T.J., and Miller, E.K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. Science *315*, 1860–1862.

Callaway, E.M. (1998). Local circuits in primary visual cortex of the macaque monkey. Annu. Rev. Neurosci. *21*, 47–74.

Cohen, M.R., and Maunsell, J.H. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nature Neuroscience *12*, 1594–1600.

Cohen, M.R., and Maunsell, J.H.R. (2010). A Neuronal Population Measure of Attention Predicts Behavioral Performance on Individual Trials. Journal of Neuroscience *30*, 15241–15253.

Cohen, M.R., and Maunsell, J.H.R. (2011). Using Neuronal Populations to Study the Mechanisms Underlying Spatial and Feature Attention. Neuron *70*, 1192–1204.

Douglas, R.J., and Martin, K.A.C. (2004). Neuronal circuits of the neocortex. Annu. Rev. Neurosci. *27*, 419–451.

Duncan, J., Ward, R., and Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. Nature *369*, 313–315.

Ecker, A.S., and Tolias, A.S. (2014). Is there signal in the noise? Nature Neuroscience *17*, 750–751.

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.

Ecker, A.S., Berens, P., Cotton, R.J., Subramaniyan, M., Denfield, G.H., Cadwell, C.R., Smirnakis, S.M., Bethge, M., and Tolias, A.S. (2014). State Dependence of Noise Correlations in Macaque Primary Visual Cortex. Neuron *82*, 235–248.

Ecker, A.S., Denfield, G.H., Bethge, M., and Tolias, A.S. (2016). On the Structure of Neuronal Population Activity under Fluctuations in Attentional State. Journal of Neuroscience *36*, 1775–1789.

Engel, T.A., Steinmetz, N.A., Gieselmann, M.A., Thiele, A., Moore, T., and Boahen, K. (2016). Selective modulation of cortical state during spatial attention. Science *354*, 1140–1144.

Eriksen, C.W., and St James, J.D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. Percept Psychophys *40*, 225–240.

Eriksen, C.W., and Yeh, Y.Y. (1985). Allocation of attention in the visual field. J Exp Psychol Hum Percept Perform *11*, 583–597.

Fitzpatrick, D., Lund, J.S., and Blasdel, G.G. (1985). Intrinsic connections of macaque striate cortex: afferent and efferent connections of lamina 4C. J. Neurosci. *5*, 3329–3349.

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

Gregoriou, G.G., Gotts, S.J., Zhou, H., and Desimone, R. (2009). High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. Science *324*, 1207–1210.

Haefner, R.M., Berkes, P., and Fiser, J. (2016). Perceptual Decision-Making as Probabilistic Inference by Neural Sampling. Neuron *90*, 649–660.

Hansen, B.J., Chelaru, M.I., and Dragoi, V. (2012). Correlated variability in laminar cortical circuits. Neuron *76*, 590–602.

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.

Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. (Lond.) *195*, 215–243.

van Kerkoerle, T., Self, M.W., and Roelfsema, P.R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. Nature Communications *8*, 13804.

Latimer, K.W., Yates, J.L., Meister, M.L.R., Huk, A.C., and Pillow, J.W. (2015). NEURONAL MODELING. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science *349*, 184–187.

Livingstone, M.S., and Hubel, D.H. (1984). Anatomy and physiology of a color system in the primate visual cortex. J. Neurosci. *4*, 309–356.

Lund, J.S. (1988). Anatomical organization of macaque monkey striate visual cortex. Annu. Rev. Neurosci. *11*, 253–288.

Lund, J.S., Lund, R.D., Hendrickson, A.E., Bunt, A.H., and Fuchs, A.F. (1975). The origin of efferent pathways from the primary visual cortex, area 17, of the macaque monkey as shown by retrograde transport of horseradish peroxidase. The Journal of Comparative Neurology *164*, 287–303.

Maunsell, J.H., and van Essen, D.C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. J. Neurosci. *3*, 2563–2586.

McAdams, C.J., and Maunsell, J.H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. The Journal of Neuroscience *19*, 431–441.

McAdams, C.J., and Reid, R.C. (2005). Attention modulates the responses of simple cells in monkey primary visual cortex. J. Neurosci. *25*, 11023–11033.

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.

Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. Physiol. Rev. *65*, 37–100.

Motter, B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. Journal of Neurophysiology *70*, 909–919.

Müller, M.M., Teder-Sälejärvi, W., and Hillyard, S.A. (1998). The time course of cortical facilitation during cued shifts of spatial attention. Nat. Neurosci. *1*, 631–634.

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.

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

Quiroga, R.Q., Nadasdy, Z., and Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput *16*, 1661–1687.

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

Rockland, K.S., and Pandya, D.N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. Brain Res. *179*, 3–20.

Roelfsema, P.R., Lamme, V.A., and Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. Nature *395*, 376–381.

Shan, K.Q., Lubenov, E.V., and Siapas, A.G. (2017). Model-based spike sorting with a mixture of drifting t-distributions.

Smith, M.A., Jia, X., Zandvakili, A., and Kohn, A. (2013). Laminar dependence of neuronal correlations in visual cortex. Journal of Neurophysiology *109*, 940–947.

Snodderly, D.M., and Gur, M. (1995). Organization of striate cortex of alert, trained monkeys (Macaca fascicularis): ongoing activity, stimulus selectivity, and widths of receptive field activating regions. J. Neurophysiol. *74*, 2100–2125.

Softky, W.R., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. The Journal of Neuroscience *13*, 334–350.

Tolias, A.S., Ecker, A.S., Siapas, A.G., Hoenselaar, A., Keliris, G.A., and Logothetis, N.K. (2007). Recording chronically from the same neurons in awake, behaving primates. J. Neurophysiol. *98*, 3780–3790.

Ungerleider, L.G., Galkin, T.W., Desimone, R., and Gattass, R. (2008). Cortical Connections of Area V4 in the Macaque. Cerebral Cortex *18*, 477–499.

Verhoef, B.-E., and Maunsell, J.H.R. (2017). Attention-related changes in correlated neuronal activity arise from normalization mechanisms. Nature Neuroscience.

Wichmann, F.A., and Hill, N.J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys *63*, 1293–1313.

Wichmann, F.A., and Hill, N.J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. Percept Psychophys *63*, 1314–1329.

Yu, B.M., Cunningham, J.P., Santhanam, G., Ryu, S.I., Shenoy, K.V., and Sahani, M. (2009). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. J. Neurophysiol. *102*, 614–635.

Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature *370*, 140–143.

Figure Legends

Figure 1. Attention and correlated variability.

A) Hypothesis 1: Attentional gain is increased, but relatively stable under both conditions (top left). Correlated variability is driven by a common noise source (top right), which is suppressed by attention (Mitchell et al. 2009, Cohen & Maunsell 2009). **B)** Hypothesis 2: Attentional gain is increased, but fluctuates from trial to trial (Cohen & Maunsell 2010, 2011, Ecker at al. 2016). Correlated variability is driven by fluctuations of attentional state. The reduction in correlations under attention would imply that the attentional gain is less variable when attending.

Figure 2. Predicted effects of attention on correlations when attending one ("Attend In/Out") or two stimuli ("Attend Both").

A) Scenario in which attentional fluctuations are negligible and attention primarily acts by suppressing common noise sources. In this case, we expect intermediate correlations when attending two stimuli. **B)** Scenario in which fluctuations in attention induce correlations. In this case, we expect attention to switch randomly between the two targets in the "Attend Both" condition, resulting in the highest correlations in this condition. **C)** Both mechanisms may contribute to different degrees. The relevance of attentional fluctuations is revealed by the difference between focused attention and split attention conditions.

Figure 3. Task diagram with behavioral results.

A) Orientation change-detection task. Two stimuli (L: left, R: right) randomly change their orientation during the ZCP (length 10-5000ms). One stimulus (R in this example) then enters the CP (300ms) when the signal orientation is shown (coherence exaggerated for clarity). This period is followed by another 200ms ZCP to allow time for a behavioral response. **B**) Illustration of attention conditions. Attention is cued according to fixation spot color. This color scheme is used in all figures to represent each condition. Percentages below the stimuli indicate the probability that the change occurs in this stimulus on a given trial. One stimulus overlaps the recorded neurons' receptive fields. **C**) Example session psychophysical performance. Individual points represent fraction of changes detected at a given coherence. Solid lines indicate fit of logistic

function to the data. Inset shows detection threshold with 95% CIs. **D**) Behavioral summary. Same as inset in **c**, but averaged across sessions in our dataset (N=27; mean±SEM). **E)** Percentage of changes detected in each condition averaged across sessions (mean±SEM).

Figure 4. Attentional modulation of neuronal responses.

A) Example session spike density function for each condition, normalized to the average response in AI condition (mean across units). **B)** Fractional increase in firing rates in the AB and AI conditions relative to the AO condition averaged across sessions (N=27; mean±SEM). **C)** Example single unit tuning curves in AI (red) and AO (blue) conditions. Dots show responses to specific orientations, solid lines show fitted von Mises functions.

Figure 5. Effects of attention on shared variability.

A) Spike count correlations from 0-1s following stimulus onset, averaged across sessions (N=27). **B**) Spike count correlations shown separately for both subjects during fixation (300ms interval) and during the task (same interval as in A). **C**) Cumulative correlation coefficient, calculated by integrating the cross-correlogram, for each attention condition and averaged across sessions. Data in A-B show mean ± SEM, C omits SEM.

Figure 6. Laminar profile of attention effects.

A) Example session CSD profile evoked by task stimulus (left column) with multi-unit receptive fields (middle) and tuning curves (right). Depths are relative to first L5 channel. Dotted black line shows L4-5 transition. Arrow shows initial current sink-source flip in L4C. **B**) Fractional increase in firing rates in AB and AI, relative to AO, conditions split by laminar group. **C**) Spike count correlation over 0-1s interval split by laminar group. Data in B-C show mean across sessions ± SEM (N=27).

640 # Materials and Methods

641

642 ## EXPERIMENTAL MODEL AND SUBJECT DETAILS

643 All behavioral and electrophysiological data were obtained from two healthy, male rhesus
644 macaque (*Macaca mulatta*) monkeys (B and D) aged 12 and 13 years and weighing 11 and 10 kg,
645 respectively, during the time of study. All experimental procedures complied with guidelines of
646 the NIH and were approved by the Baylor College of Medicine Institutional Animal Care and
647 Use Committee (permit number: AN-4367). Animals were housed individually in a large room
648 located adjacent to the training facility, along with around ten other monkeys permitting rich
649 visual, olfactory and auditory interactions, on a 12h light/dark cycle. Regular veterinary care and
650 monitoring, balanced nutrition and environmental enrichment were provided by the Center for
651 Comparative Medicine of Baylor College of Medicine. Surgical procedures on monkeys were
652 conducted under general anesthesia following standard aseptic techniques. To ameliorate pain
653 after surgery, analgesics were given for 7 days. Animals were not sacrificed after the experiments.

654

655 ## METHOD DETAILS

656 ### Visual stimuli and behavioral paradigm

657 Visual stimuli were two Gabor patches (size: 2–3° depending on eccentricity; spatial frequency:
658 3–3.5 cycles per degree; contrast: 100% Michelson) presented on CRT monitors (at a distance of
659 100 cm; resolution: 1600 × 1200 pixels; refresh rate: 100 Hz) using Psychophysics Toolbox
660 (Brainard, 1997). The monitors were gamma corrected to have a linear luminance response
661 profile. Video cameras (DALSA genie HM640; frame rate 200Hz) with custom video eye tracking
662 software developed in LabView were used to monitor eye movements.

663 Monkeys performed a noisy, orientation–change detection task. Trials were initiated by a
664 sound and the appearance of a colored fixation target (~0.15°). Monkeys were required to fixate
665 within a radius of 0.5°–1°, but typically fixated much more accurately, as revealed by offline
666 analysis. After fixating for 300ms, two Gabor patches were presented symmetrically in the lower
667 left and right visual fields. During what we labeled the Zero-Coherence Period (ZCP), these
668 stimuli changed their orientation pseudo-randomly every 10ms (uniform distribution over 36

orientations spaced by 5° between 0 and 175°) for a random period of time drawn from an exponential distribution with a minimum of 10ms, mean of 2170ms, and maximum of 5000ms.

After this time one of the two stimuli entered the Coherent Period (CP), where one particular orientation, called the "signal" orientation, was shown with a higher frequency than the other orientations. The CP lasted 300ms (30 frames), and from trial to trial the number of frames in the CP showing the signal orientation was selected from a set of five unique "coherences" chosen for that session, which allowed us to vary the difficulty of the trials within a session and compute psychometric functions. After this period, the stimulus returned to the ZCP for a further 200ms to allow sufficient time for subjects to report whether or not they noticed the presence of the signal orientation by making a saccade to the stimulus showing the change. Subjects were prevented from responding within the first 100ms of the CP to minimize guessing. Successful identification of the signal orientation was rewarded with a small drop of juice. On 10% of trials in each attention condition no change occurred, and subjects were rewarded for maintaining fixation. Orthogonal signal orientations were used in the left (135°) and right (45°) stimuli.

Note, occurrences of the signal orientation during the CP were not constrained to occur in successive frames. Also note that the left and right stimuli displayed different orientation sequences, so that subjects could not identify a change simply by noticing when the two orientation sequences diverged. Orientation sequences were described as pseudo-random for the following reason. For each trial a random number generator seed was chosen from a set of five such seeds selected for a given recording session. Doing so meant there were five unique stimuli that could be repeated across attention conditions for the purposes of calculating spike count correlations and Fano factors over identical stimuli. Sequences were constrained to show each orientation once before any repetitions were allowed so that the maximum number of signal orientations that could occur by chance in a period of time equal to the CP (300ms) was two.

Attention was cued in blocks of trials by the color of the fixation spot (Fig. 3B). In the Attend Out (AO) condition, 100% of the changes occurred in the non-receptive field stimulus. In the Attend In (AI) condition, 100% of changes occurred in the receptive field stimulus. In the Attend Both (AB) condition, the change was equally likely to occur in either stimulus (50% chance that the change was in the receptive field stimulus). Block transitions occurred after a total of 60 hit

698 and miss trials was achieved (i.e. false alarms did not count). Blocks were randomized in sets of
699 three so that each attention condition was seen before one was allowed to repeat. Coherences
700 were increased by one frame in the AB condition to keep task difficulty approximately constant
701 across conditions.
702

703 **Surgical methods**

704 Our surgical procedures followed a previously established approach (Tolias et al., 2007). A cranial
705 headpost was first implanted under general anesthesia using aseptic conditions in a dedicated
706 operating room. After premedication with Dexamethasone (0.25–0.5 mg/kg; 48 h, 24 h and on the
707 day of the procedure) and atropine (0.05 mg/kg prior to sedation), animals were sedated with a
708 mixture of ketamine (10 mg/kg) and xylazine (0.5 mg/kg). During the surgery anesthesia was
709 maintained using isoflurane (0.5–2%).

710 After subjects were trained to perform the above described task, they were implanted with a
711 form-fitted titanium recording chamber, designed based on pre-operatively obtained anatomical
712 MRI scans, placed at a location over the operculum in V1 determined by stereotactic coordinates
713 (Tolias et al., 2007). This surgery was performed under identical conditions as described for
714 headpost implantation. The chamber was attached to the skull using orthopedic screws only. We
715 used a small amount of dental cement to seal any openings between the bone and the lower
716 surface of the recording chamber. A custom-made chamber cap was then placed to seal the
717 chamber and prevent infection. A minimum of three weeks was provided for the implant to heal.
718 After healing, small 2–3mm trephinations could be performed, in aseptic conditions under
719 ketamine (10 mg/kg) sedation with ketoprophen (2mg/kg) for analgesia and meloxicam
720 (0.2mg/kg for two days), to enable access for subsequent daily electrophysiological recordings.
721

722 **Electrophysiology in awake, behaving monkeys**

723 We performed daily electrophysiological recordings beginning 48 hours after a craniotomy was
724 performed. Custom-designed 32 channel, linear silicon probes (NeuroNexus V1x32-Edge-10mm-
725 60-177) were mounted in a Narishige microdrive (MO-97) with a nested, stainless steel guide tube
726 composed of one extra-thin walled 23-gauge piece, spanning most of the length of the probe shaft,

727 and a smaller 27-gauge piece (roughly 6mm long) nested inside such that 4mm of the smaller
728 tubing protruded beyond the large piece. This design enabled a tight fit around the probe to
729 support it during dural penetrations. We took care during the insertion procedure to ensure that
730 the dura was penetrated only by the probe itself, rather than the guide tube, to minimize damage
731 to the superficial layers of cortex. We alternated lowering the guide tube in steps of 250μm and
732 extending the probe up to ~500μm beyond the guide tube, retracting and repeating as necessary,
733 until either characteristic changes in the LFP or multi-unit activity, or both, were observed,
734 indicating successful penetration of cortex.

735 The probe was then lowered in ~250μm steps at < 10μm per second, pausing for several
736 minutes after each step, until activity was seen on all channels. As a result of this procedure there
737 would be variable degrees of tissue compression. Some of this compression was relieved early in
738 the positioning of the probe by retracting the guide tube by ~500μm after the probe was several
739 hundred microns inside the cortex. If compression remained after completely lowering the probe,
740 we could successfully relieve it by slowly retracting the guide tube further. The single most
741 reliable indicator of the position of our probe in cortex before receptive field mapping was a band
742 of high spontaneous activity corresponding to layer 4C (Snodderly and Gur, 1995), which could
743 be clearly seen to span roughly 6–7 channels. In general, we found the basic laminar properties
744 described by Snodderly and Gur (1995) to be very reliable guidelines. After final positioning of
745 the probe, we allowed between 30–60min for tissue settling and recording stability to become
746 established. The entire insertion procedure typically took around 3-4 hours, from penetrating the
747 dura to the start of recording. Receptive field mapping experiments were performed (see **Data**
748 **Analysis** below for details) to determine where to place one of the two stimuli such that it covered
749 the recorded neurons' receptive fields for that session.

750

751 **Data acquisition and spike sorting**

752 The methods described below for spike detection and spike sorting were adapted for use with
753 multi-channel silicon probes from our previous methods used for tetrode recordings (see Ecker
754 et al., 2014). Neural signals were digitized at 24 bits using analog acquisition cards with 30 dB of
755 onboard gain (PXI-4498, National Instruments, Austin, TX) and recorded continuously at 32 KHz

as broad-band signal (0.5 Hz to 16 kHz). Eye movement traces were sampled at 2kHz.

Spikes were detected offline when the signal on a given channel crossed a threshold of five times the standard deviation of the corresponding channel. To avoid artificial inflation of the threshold in the presence of a large number of high amplitude spikes, we used a robust estimator of the standard deviation, given by $\sigma = \text{median}(|x|)/0.6745$ (Quiroga et al., 2004). Spikes were aligned to the center of mass of the continuous waveform segment above half the peak amplitude. Code for spike detection is available online at https://github.com/atlab/spikedetection.

Virtual electrodes consisting of six channels were constructed in a sliding window (stride 2) spanning the length of the probe to aid in the spike sorting process by enabling some degree of triangulation, as with tetrodes. Given a channel spacing of 60μm, in many cases the waveforms of a single neuron could be detected by several channels. To extract features for spike sorting, we performed principal component analysis on the extracted waveform segments (individually for each channel). This step reduced the data to three dimensions per channel, resulting in an 18-dimensional feature vector. We fit a mixture of t distributions with a Kalman filter on the cluster means to track waveform drift (Shan et al., 2017).

The number of clusters was determined based on a penalized average likelihood, where the penalty term was a constant cost per additional cluster. Code for spike sorting is available online at https://github.com/aecker/moksm. Following this automatic step, results of the model were examined manually for each virtual electrode and single units were flagged at this time according to degree of cluster isolation, uniqueness of waveforms and size of refractory period. To avoid duplicate single units due to overlapping channel groups used for spike sorting, we included only those single units that had their largest waveform amplitude on one of the two central channels of the group (this was not an issue for the first and last two channels on the probe).

Dataset and inclusion criteria

Our dataset included 27 sessions (N=7, Subject B; N=20, Subject D), yielding 416 single units (N=83, Subject B; N=333, Subject D). We included recording sessions with at least 10 single units that were visually responsive and significantly orientation tuned in each attention condition. To ensure reliable estimates of neuronal (co-)variability, sessions were also excluded if there were

fewer than three (of five possible) valid seed conditions. A seed condition was considered invalid if in any of the three attention conditions there were fewer than three correct trials generated using that seed that had sufficient ZCP length available for spike count analysis. On average for the 1-second analysis window, included sessions had ~10 correct trials per seed per attention condition.

After having collected a complete dataset of 13 sessions from Subject B and a dataset of 29 sessions from monkey D, we found that sessions with recording locations close to the vertical meridian did not exhibit our predicted main effect. We reasoned that this lack of effect was likely because the two stimuli were too close to each other, allowing the monkey to attend to both simultaneously. To verify that this result was not a false positive due to post-hoc analysis, we collected an independent 10-session dataset at high eccentricities from Subject D (the termination condition of 10 sessions was set before starting to collect additional data), which confirmed the effect at high eccentricity. The results reported in this paper include all sessions with x-axis receptive field eccentricities of at least 3° in Subject B and 3.2° in Subject D (representing the median such eccentricities for each subject), including the separate validation dataset from monkey D.

Data analysis

Analysis of behavioral results

Trial results were classified as 'hits', 'misses', 'correct rejections' (for successful completion of trials with no change) and 'false alarms' (for saccades made to a stimulus before any change occurred). For each session, behavior was analyzed by calculating the fraction of changes detected (hits / [hits + misses]), both conditioned on and marginalized over coherence in each attention condition. Psychometric functions were plotted as the fraction of changes detected versus coherence in each attention condition. Using the psignifit toolbox (Wichmann and Hill, 2001a, 2001b) in MATLAB, logistic functions were fit to the attention condition specific curves using the method of maximum likelihood, and 50% performance thresholds were extracted.

Analysis of receptive fields

Prior to starting the main task, we quantitatively mapped receptive fields based on unsorted multi-unit responses using a white noise random dot stimulus. A single square dot of size 0.29 degrees of visual angle was presented on a uniform gray background, changing location and color (black or white) randomly every three frames, or 30ms, for 1 second. Receptive field profiles were obtained by spike-triggered averaging.

Analysis of orientation tuning

Our task allowed us to compute orientation tuning curves for each neuron. We binned the spike counts in bins of 10ms and used linear regression based on a one-hot encoding of the 15 stimuli directly preceding the response (i.e. the stimulus is a 36×15-dimensional vector, because there were 36 possible stimulus orientations). We defined the optimal latency of each neuron as the time delay that produced the strongest response modulation across orientations (determined by taking the variance of the regression weights across orientations). The optimal latency of most neurons was 50ms. We then re-estimated the regression using only that single time lag to obtain a tuning curve. Significance of tuning was then tested by projecting the weight vector onto a complex exponential with one cycle, the norm of which was compared to its null distribution calculated by randomly shuffling orientation labels. A p-value was obtained by performing 1,000 iterations of the shuffling procedure and using the fraction of runs in which the norm of the shuffled projection was greater than that observed in the real data. Signal correlations were computed for pairs of neurons by calculating the correlation coefficient between the two cells' tuning curves.

Analysis of gain versus offset modulation

For each unit, a von Mises distribution function, parameterized as

$$Y = w_1 + \exp(w_2 + w_3 \cos(x - w_4)),$$

was fit to the tuning curve obtained across all trials via the method described above. From this fit, the shape and preferred orientation parameters, w_3 and w_4, were obtained. These parameters were assumed not to change across attention conditions, leaving only the offset, w_1, and gain, $\exp(w_2)$,

842 terms to vary across conditions. New von Mises functions were then fit for each attention condition
843 using a linear regression model with a binary indicator variable for attention condition and an
844 interaction term. To illustrate, we write the response y to orientation i as

845 $$y_i = w_1 + \exp(w_2 + w_3 \cos(x_i - w_4)) = b_1 + b_2\theta_i$$

846 where $\theta_i = \exp(w_3 \cos(x - w_4))$ and was obtained from the overall tuning curve as described.
847 Our linear regression model comparing fits in the AO and AI condition, for example, then
848 became:

849 $$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} X_{i2}$$

850 where $X_{i1} = \theta_i$ and $X_{i2} \in \{0, 1\}$, with 0 coding the AO condition and 1 coding the AI condition.
851 In this manner we enabled different gain and offset terms to be fit to different attention conditions.
852 We then assessed whether significant attentional modulation was present by performing an F-test
853 comparing the full model above to the reduced model containing only the β_0 and β_1 terms, and
854 when significant, we tested whether the offset and gain parameters differed between conditions
855 with t-tests.
856

857 *Analysis of firing rates*

858 Visual responsiveness of neurons was determined by comparing firing rates in the 300ms fixation
859 interval before stimulus onset to those in the 300ms immediately following stimulus onset. A t-
860 test was performed to test for a significant change in rate following stimulus onset. Spike density
861 functions (SDFs) were calculated first for a given neuron, across all hit trials grouped by attention
862 condition and stimulus seed, by counting spikes in 50ms bins relative to stimulus onset and
863 averaging across trials. Averages were then taken across seeds and smoothed with a Gaussian
864 window. To calculate SDFs for a given session, individual neuron SDFs were normalized by the
865 average response in the AO condition, starting from 100ms after stimulus onset, before averaging
866 across neurons. Fractional firing rate increases were also calculated first at the individual
867 neuronal level, by averaging all available bins from the first second following stimulus onset
868 conditioned on the stimulus seed for each attention condition, and then averaging across seeds.

The rates were again normalized by the AO condition rate before averaging across neurons to get a session-level rate modulation for each attention condition. Finally, responses in the AI and AB conditions were converted to fractional changes relative to the AO responses.

Analysis of neuronal (co-)variability

Fano factors and spike count correlations were computed on the first 1000ms of the response. Fano factors were computed as the variance of the spike count divided by its mean. Spike count correlations were computed as the covariance of the two neurons' z-scored responses to identical repetitions of the same stimulus condition (seed). Z-scoring and Fano factor calculations were performed in a block-wise fashion to control for slow fluctuations in firing rate across a recording session. For the analysis of correlation timescale we used the relationship between spike count correlations and cross-correlation functions first described in Bair et al. (2001) to compute a cumulative correlation coefficient, r_{CCG}. We compute a spike train cross-correlation function for a pair of neurons *j* and *k*, as well as a shift-predictor, which is the cross-correlation function of the spike density functions of neurons *j* and *k*. The shift-predictor is subtracted from the cross-correlation function to control for stimulus-induced correlation. This shift-corrected cross-correlation is denoted $C_{jk}(\tau)$. The cumulative cross-correlation is given by

$$A_{jk} = \int_{-\tau}^{\tau} C_{jk}(t)dt$$

Following Ecker et al. (2014), the cumulative correlation coefficient is

$$r_{CCG}(\tau) = \frac{A_{jk}(\tau)}{\sqrt{A_{jj}(T)A_{kk}(T)}}$$

where T is the last time point in the counting window, in our case 1000ms.

Analysis of micro-saccades

We identified micro-saccades our subjects made during the ZCP of our task (when spike counts

were analyzed) to determine whether our correlation results could be accounted for by an increase in micro-saccade frequency in our AB condition, relative to the AI and AO conditions. We followed the definitions described in Bair and O'Keefe (1998). Periods of stable gaze were taken to be those intervals during which eye position remained within a 0.1-degree window. Deviations greater than 0.1 degree in 10ms (10deg/s velocity) were taken to be micro-saccades. The number of micro-saccades during analysis periods was counted for each attention condition in each session and a two-factor ANOVA was performed to determine whether micro-saccades differed across conditions.

Analysis of laminar data

The CSD profile at each time point was calculated following Mitzdorf (1985) as the second spatial derivative of the task-stimulus evoked LFPs across channels, smoothed with a Gaussian kernel to aid visualization. The granular layer was identified according to several criteria used in conjunction. The earliest current sink to source transition (identified by an arrow in Fig. 6A) is one indicator, immediately below which is a complementary source to sink transition in L5. We used additional criteria, described by Snodderly and Gur (1995), to verify this positioning, because there was a prominent current sink to source transition in L6 as well. These criteria included higher spontaneous activity and more poorly defined orientation tuning curves characteristic of the granular layer (Snodderly and Gur, 1995). Additional reports have described the granular layer to contain smaller receptive fields (Hubel and Wiesel, 1968; Livingstone and Hubel, 1984), which we also saw (Fig. 6A). In general across sessions, all of these granular layer features were quite consistent, allowing for confident determination of the L4-5 boundary. The first L5 channel was labeled as the zero-point for depth. Negative depths are more superficial to this point. The granular layer was defined as a roughly 400μm band just superficial to the zero-point (Fitzpatrick et al., 1985; Hansen et al., 2012; Lund, 1988; Smith et al., 2013). The supragranular group (L1–3) was defined as everything superficial to the top of the granular layer, and the infragranular group (L5–6) was defined as everything deeper than and including the zero-point.

QUANTIFICATION AND STATISTICAL ANALYSIS

Although customary in the field, we did not consider units or pairs as independent samples. Treating units as independent samples ignores the session-to-session variability and leads to underestimated confidence intervals and, consequently, inflated false positive rates. Instead, we first averaged our measurements across observations within a session and then performed all statistical tests across sessions, treating the session averages as independent samples. While this approach sacrifices some statistical power, it leads to conservative estimates of p values.

For statistical analyses involving our attention conditions, two-factor ANOVAs were used, with session and attention condition as the two factors. The Tukey-Kramer method was used for post-hoc analyses. The only exception is the test for significantly elevated AB condition correlations, where we performed a one-tailed t-test on a contrast between the AB condition and the average of the AO and AI condition results. This choice is justified by our previously published model (Ecker et al., 2016), which predicts this effect and its direction and was hypothesized and specified before data collection. For assessments of visual responsiveness and significant increases in fractional firing rates, two-tailed t-tests were used, which, for rate increases, were Bonferroni-corrected for multiple comparisons. Orientation tuning significance was assessed according to the permutation test described above. Statistical comparisons were considered significant at $p < 0.05$ ($p < 0.0167$ for Bonferroni-corrected tests for firing rates in association with Figure 4B, as there were 3 comparisons; $p < 0.025$ for those associated with Figure 6B, given two comparisons). All error bars show the standard error of the mean (SEM; either directly calculated or estimated via two-factor ANOVA), except in the Figure 3C inset, which shows 95% confidence intervals.

Data availability

The datasets generated during and analyzed during the current study, along with the code to replicate the presented analyses, are available from the corresponding author on reasonable request.