



HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2010 June 01.

Published in final edited form as:

Nat Neurosci. 2009 December ; 12(12): 1594–1600. doi:10.1038/nn.2439.

Attention improves performance primarily by reducing interneuronal correlations

Marlene R Cohen and John HR Maunsell

Howard Hughes Medical Institute and Harvard Medical School Department of Neurobiology, Boston, MA, 02155

Abstract

Visual attention can dramatically improve behavioural performance by allowing observers to focus on the important information in a complex scene. Attention also typically increases the firing rates of cortical sensory neurons. Rate increases improve the signal-to-noise ratio of individual neurons, and this improvement has been assumed to underlie attention-related improvements in behaviour. We recorded dozens of neurons simultaneously in visual area V4 and found that changes in single neurons accounted for only a small fraction of the improvement in the sensitivity of the population. Instead, over 80% of the attentional improvement in the population signal was caused by decreases in the correlations between the trial-to-trial fluctuations in the responses of pairs of neurons. These results suggest that the representation of sensory information in populations of neurons and the way attention affects the sensitivity of the population may only be understood by considering the interactions between neurons.

Introduction

The responses of sensory neurons are variable, and laboratory studies typically deal with this variability by averaging responses to many stimulus presentations. In the real world, however, people and animals must respond to individual stimulus events, and the brain is thought to compensate for neuronal variability by encoding sensory information in the responses of large populations of neurons. To understand the way sensory information guides behavior in everyday life, we need to understand the way information is encoded in populations of neurons.

One way to identify the important aspects of a population code is to look at the differences between the neuronal representation of a sensory stimulus when it is used to guide behavior and when it is behaviourally irrelevant. Tasks that control attention provide a powerful way to manipulate behavioural relevance. Attention allows observers to select the most important stimuli and greatly improves perception of the attended location or feature. Attention

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

Correspondence and requests for materials should be addressed to M.R.C. (marlene_cohen@hms.harvard.edu).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author information M.R.C. collected the data and performed the analyses. M.R.C. and J.H.R.M. designed the study and wrote the paper. The authors have no competing financial interests.

modulates the firing rates of sensory neurons, typically increasing responses to attended stimuli^{1–4}. This increased rate of firing acts to improve the signal-to-noise ratio of individual neurons^{5, 6}, and a recent study found that attention can cause a small additional reduction in the mean-normalized variance (Fano factor) of the responses of some neurons in visual area V4⁷. However, the net effect of attention on the signal-to-noise ratio of single neurons is modest, suggesting that attention causes large improvements in psychophysical performance by affecting population responses in ways that cannot be measured in single neurons.

Attention could also alter the reliability of neuronal representations by affecting the amount of noise that is shared across a population of neurons. Variability in a population depends in part on the variability of single neurons, but can depend greatly on the extent to which variability is shared across the population. The effect of correlated variability on population sensitivity depends on the way in which the population is read out^{8, 9}, but its effect can be far greater than the effect of independent variability of single neurons. If the noise in individual neurons is independent, averaging the responses of many neurons will lead to a very accurate estimate of the mean, no matter how noisy the individual neurons are. If, however, there are positive correlations in the trial-to-trial fluctuations of the responses of pairs of neurons, then the shared variability can never be averaged out, leading to a more variable (and less accurate) estimate of the mean activity in the population^{10–12}.

Results

We investigated the effect of attention on both single neuron responses and correlated variability by recording from populations of neurons in visual area V4 using chronically implanted microelectrode arrays in two rhesus monkeys (*Macaca mulatta*). Each animal had two arrays, allowing us to monitor populations of neurons in both hemispheres simultaneously. Figure 1A shows the centers of the multiunit receptive fields we recorded from one of our monkeys. The diameter of V4 receptive fields is approximately equal to eccentricity^{13, 14}, so the receptive fields the neurons recorded in a hemisphere typically overlapped at least partially. We recorded from 376 single units and 2746 multiunit clusters during 41 days of recording (including 66,578 simultaneously recorded pairs in the same hemisphere and 59,990 pairs in opposite hemispheres). We did not find any important differences between single and multiunits or between the two monkeys and our population analyses require large neural populations, so we combined single and multiunits here (see Supplementary information). However, the statistics in the text for single units are based on a subset of 187 single units that we are confident are unique (if there was a single unit on a given electrode on multiple days, it was only counted once).

The monkeys performed an orientation change detection task in which spatial attention was manipulated (Fig 1B). Two Gabor stimuli flashed on and off and the monkey's task was to detect a change in the orientation of either stimulus. We manipulated attention in blocks by cueing the monkey as to which stimulus was more likely to change (see Methods). Each day, the location, size, orientation, and spatial frequency of the Gabors were optimized for a selected single unit in each hemisphere. The two stimuli were therefore different, so directing attention to one of the two stimuli likely modulated feature-based as well as spatial

attention. Because we recorded from neurons with a wide range of receptive field locations and tuning, most neurons were not well driven by the stimulus (mean driven rate was 8.2 spikes/s for single units and 21.5 spikes/s for multiunits compared to mean spontaneous rate 5.8 spikes/s for single units and 14.1 spikes/s for multiunits).

Attention greatly improved behavioural performance in this task. To motivate the monkeys to attend to the cued location, the stimulus at the attended location was the one that changed on 80% of trials (trials in which the attentional cue was valid). On the remaining 20% of trials (invalid trials), we tested performance at the unattended location using only a single orientation change (11°), which allowed us obtain reliable estimates of behavior and neural responses even given the relatively few invalid trials. Figure 1C shows psychometric data from a typical recording session, in which the proportion of trials on which the monkey successfully detected an 11° orientation change was substantially greater on trials when the attended (black) rather than the unattended (grey) stimulus changed orientation.

To compute the effects of attention on neural responses during the period in which the monkey's attentional state was most likely to affect its behavioural performance, we focused most analyses on the stimulus presentation directly preceding the orientation change (black outlined box in Fig 1B). On a given day, the stimuli immediately before the orientation change were identical, regardless of the attentional condition, validity of the attentional cue, or size of the orientation change. Invalid trials were randomly interleaved with valid trials, so the neuronal effects of attention were indistinguishable on valid and invalid trials. We observed some adaptation of V4 responses between the first and the second stimulus presentation on each trial, but the average responses to the second through tenth stimuli were statistically indistinguishable (t-tests, $p > 0.5$). Because the orientation change occurred no sooner than the third stimulus presentation, the responses to the stimulus directly before the change was unaffected by the length of the trial.

Consistent with previous studies^{3, 15-18}, we found that attention increased V4 firing rates (Figure 2A). To quantify the increase, we calculated a standard modulation index (MI_{rates}), which was the difference between the average firing rates on trials when the attended stimulus was inside or outside the neuron's receptive field trials divided by the sum (see Supplementary Material). The mean MI_{rates} was 0.049 for single units and 0.042 for multiunits, both of which were significantly greater than zero (t-tests, $p < 10^{-6}$ for single units, $p < 10^{-20}$ for multiunits).

We also found that attention reduces the trial-to-trial variability of individual neurons over a similar time course to its effects on firing rate. As is common to stimulus responses in many cortical areas¹⁹, we observed a drop in the Fano factor (the ratio of the variance of the firing rates to the mean) following stimulus onset (Fig. 2B). Following the drop associated with the response transient, the Fano factor remains at a significantly lower level in the attended than in the unattended condition (mean MIFF during the sustained response was -0.011 for single units and -0.017 for multiunits, $p < 0.05$ and $p < 10^{-3}$, respectively). Because the Fano factor plotted in Figure 2B was calculated using subdistributions of neurons such that the mean firing rates were the same for each time point and attentional condition^{19, 20}, the time

course and attentional-dependence of Fano factor are independent of changes in firing rate (see Supplementary Material).

The data in Figures 2A and 2B show that attention improves the signal-to-noise ratio of individual V4 neurons, but we found that the effect of attention on the correlated variability in pairs of neurons was even more important. For each pair of simultaneously recorded neurons and each attentional condition, we calculated the correlation coefficient between spike count responses to the stimulus preceding the orientation change.

This metric, termed noise correlation, measures the correlation in trial-to-trial fluctuations in responses, and therefore has a very different timescale than the millisecond timescale synchrony that has been shown to increase with attention²¹. We did not focus on synchrony here because no more pairs than expected by chance exhibited significant synchrony (3609 significant pairs out of 66,578 pairs = 5.4% in the attended condition and 3634 significant pairs in the unattended condition, 5.5% $p < 0.05$, bootstrap test described in Methods) and synchrony in the attended and unattended conditions were not different (paired t-test, $p = 0.46$). Many spikes are needed to detect statistically significant synchrony, and even more to detect modulation of synchrony by processes such as attention. Synchrony has therefore been observed in some studies of visual cortex (see for example^{21, 22}) but not others^{10, 23, 24} (see²¹ for a discussion of the statistical power needed to detect synchrony). The absence of synchrony in our study is likely due to a combination of the low firing rates of many of our cells caused by stimuli that were suboptimal for most cells, the fact that we calculated synchrony using pairs of spiking neurons rather than correlating spike times with local field potentials (see²¹), and the fact that most neuron pairs were separated by millimeters in the cortex. The correlations we observed were fluctuations on a longer timescale than millisecond-level synchrony. One possibility is that the same mechanisms that cause low frequency oscillations in EEGs and local field potentials (which have been shown to desynchronize with attention^{25–28}) cause the correlations we measured.

To obtain accurate estimates of noise correlation, we did not calculate a time course of correlation as we did for rate and Fano factor, because over short periods, the distributions of spike counts become non-Gaussian (because spike counts can never be negative) and discrete. Skewed, discrete distributions pose a problem for second-order statistics like correlation, causing noise correlations to approach zero as the mean number of spikes decreases^{22, 29, 30}. We therefore calculated noise correlation over the entire 200 ms interval (Figure 2C).

Because the stimuli produced a wide range of responses across the population of neurons, we binned the neuron pairs by their mean evoked response across both attentional conditions (driven rate – baseline) in Figure 2C. Noise correlations were highest for pairs of neurons in the same hemisphere (solid lines) that both responded strongly to the stimulus. This result can be explained by the fact that correlations tend to increase with firing rate³⁰ and the observation that noise correlations are highest for neurons with similar tuning^{10, 22, 24, 29}. Our data set included neurons with a broad range of preferences for orientation and other stimulus properties and different receptive field locations, so two neurons that were both strongly modulated by the stimulus likely had similar tuning.

To test the effect of tuning similarity on noise correlation more directly, we calculated noise correlation as a function of signal correlation (Figure 2D). In a separate set of trials, we presented Gabor stimuli at a variety of locations and orientations while the monkey performed a change detection task far outside the neurons' receptive fields (see Methods). We calculated signal correlation by computing a correlation between the mean responses of each neuron to each stimulus. Consistent with previous results^{10, 22–24}, we found that noise correlation is highest for neurons with similar tuning (large, positive signal correlation) and lowest for neurons with opposite tuning (negative signal correlation). Unlike a recent study of noise correlations in V1 using the same electrode arrays as we used here²⁹, we found that noise correlation did not depend on cortical distance. We suspect that the greater retinotopic and tuning organization of V1 compared to V4 accounts for the differences in our results.

We found that even for the least responsive neurons (Figure 2C) and pairs of neurons with dissimilar stimulus preferences (Figure 2D), correlations within a hemisphere were on average positive, indicating that there is shared variability throughout the population. In contrast, we found that noise correlations for pairs of neurons in opposite hemispheres were close to zero, meaning that within an attentional condition, trial-to-trial fluctuations in the two hemispheres are independent.

The biggest physiological effect of attention in our data set was a large decrease in the correlations between pairs of neurons in the same hemisphere (compare the black and grey solid lines in Figures 2C and 2D). On average, attention reduced noise correlations by about half (mean $MI_{\text{cor}} = -0.35$ for single units and -0.29 for multiunits, $p < 10^{-5}$ for single units and $p < 10^{-9}$ for multiunits). Attentional modulation of correlation depended strongly on how much the neurons were driven by the stimulus: for the most responsive pairs of neurons, noise correlation in the attended condition was roughly one-third the correlation in the unattended condition (Figure 2C). In contrast, the effect of attention on correlations did not depend on the degree of tuning similarity between the two cells (the black and grey lines in Figure 2D are parallel). This observed decrease in correlation due to attention is the opposite result predicted by the mathematical relationship between firing rate and correlation³⁰. Attention tends to increase firing rates (Figure 2A) which makes the distributions of spike counts more Gaussian and less discretized, leading to a predicted *increase* in correlation. Therefore, decreases in correlation cannot be a simple mathematical consequence of increases in firing rate.

Previous studies have shown that attention modulates firing rate more for neurons with the biggest response to the stimulus (whose responses may be more informative for the task)^{15–18}. Furthermore, a recent study found that fast spiking neurons with high firing rates (putative interneurons, separated from putative excitatory neurons on the basis of waveform width) showed larger differences in Fano factor than regular spiking neurons⁷. Our hardware filters prevented us from distinguishing these neuron types on the basis of waveform, but consistent with this study and studies of attentional modulation of firing rate, we found that attention has a bigger effect on the rates, Fano factors and noise correlations of neurons that responded strongly to the stimulus. Figures 3A–C plot the difference in firing rate, Fano factor, and correlation between the two attentional conditions (attended – unattended). Neurons (or pairs of neurons) that were most strongly driven by the stimulus (biggest

difference between evoked and baseline firing rate) likely have receptive field locations and tuning properties that make them well suited for this task, and these neurons show the largest effects of attention by all three measures.

Recording from both hemispheres simultaneously allows us to be sure that the correlation changes we observed are spatially-specific effects of attention. The same block of trials that yielded low correlations in one hemisphere gave high correlations in the other, so non-specific factors such as arousal or motivation cannot account for the changes in correlation we observed. The fact that trial-to-trial variability in the two hemispheres was virtually independent is further evidence that the correlation changes we observed within a hemisphere are spatially specific.

Consistent with many previous studies (for examples, see ^{21, 31-34}), we found that attention primarily affects the sustained part of the response rather than the onset transient (see timecourses in Figures 2A and 2B). In our data, attentional modulation of firing rate becomes statistically significant 122 ms following stimulus onset (first time point at which the 95% confidence intervals for the means of the two attentional conditions do not overlap). In addition to examining the effect of attention on rates, Fano factor, and correlations during the entire stimulus period, we calculated attentional effects for all three measures during the sustained response of the response (122 ms to 260 ms following stimulus onset). As expected, attentional effects were larger during the sustained period by what appeared to be a fairly constant factor (the dashed and solid lines in Figure 3 are approximately scaled versions of each other).

The data in Figures 2 and 3 show that attention changes the responses of both single neurons and correlated variability in ways that could allow each to contribute to improvements in population sensitivity. A primary goal of this study was to determine the relative importance of changes in firing rates, Fano factor, and noise correlations. Because Fano factor measures the variability of single neurons without regard to the source of that variability, the decrease in Fano factor that we observed (Figures 2B and 3B) could arise from a decrease in the independent variability of individual neurons, a decrease in shared variability across the population, or a combination of both. Noise correlation measures the degree of shared variability. We therefore focused on the other aspect of variability captured by the Fano factor, asking how much a decrease in independent variability that was large enough to account for the full decrease in Fano factor would improve population sensitivity.

Using the procedure described in Figure 4, we compared the effects of attentional modulation of firing rates, independent variability, and noise correlation, and we found that the modulation of correlation has by far the greatest influence on the attentional improvement in population sensitivity. We first quantified the degree to which attention improved the sensitivity of the groups of neurons we recorded (Figures 4A and 4C), then determined the amount of that improvement that was caused by attentional modulation of that affected only rate, only independent variability, or only correlation (Figures 4D and 4E).

We quantified how much attention improved neuronal signals in our recorded populations using an approach that is schematized for a hypothetical two-neuron data set in Figure 4A

and shown for a real 38-neuron data set in Figure 4C. The monkey's task was to detect a change in the orientation of the stimulus, so we defined population sensitivity as the discriminability between the distributions of responses to the original orientation and the changed orientation. For each of the single and multiunits we recorded from a given hemisphere on a given day (mean 39.5 neurons, range 14 to 74), we calculated responses to the stimulus preceding the change from 60 ms to 260 ms following stimulus onset and the changed stimulus starting 60 ms after onset and continuing for either 200 ms or until 60 ms before the animal's response, whichever came first. The mean time from the onset of the changed stimulus to the onset of the animal's response was 251 ms, and 260 ms fell at least 60 ms before the saccade on 39% of trials. We experimented with other intervals for computing spike counts (including identical periods for the original and changed stimuli (200 ms each) and also cutting off the response to the changed stimulus 100 ms or 0 ms before the saccade), and these did not qualitatively affect the proportion of the improvement in population sensitivity accounted for by each of the three factors we considered. Using this time period, attentional modulation during the changed stimulus was indistinguishable from modulation during the previous stimulus (see Supplementary Figure 1).

We plotted one point for each stimulus in each trial in an n -dimensional space in which each dimension corresponds to the response of one of the n neurons we recorded in a given hemisphere (Figure 4A). We then calculated the mean response for each stimulus and projected all responses onto an "axis of discrimination" drawn through the two means. This was done separately for the two attention conditions, producing pairs of one-dimensional distributions of projections for each attention condition (see the left column of Figure 4C for these distributions from an example data set).

We measured population sensitivity by calculating d' (the difference in the means divided by their RMS standard deviation), which is monotonically related to theoretical performance on classifying stimuli, so attention should increase d' to improve behavioural performance. We quantified the attentional improvement in population sensitivity as the difference in d' between the attended and unattended conditions. The placements of the distributions on the y-axis in Figure 4C correspond to the measured d' (left axis), so the higher d' in the attended condition places these distributions above the lower d' in the unattended condition. We then normalized the d' values to reflect the measured improvement (right axis).

The amount of attentional improvement in our d' measure correlates strongly with the monkey's behavioural improvement due to attention. For each hemisphere-day, we quantified behavioural improvement as the lateral shift between measured performance in the unattended condition and the fitted psychometric curve in the attended condition (Figure 1C and Methods). In the example in Figure 1C, attention shifted the psychometric curve by 7.7° , which was typical for our data sets (mean = 7.6° , 0.5° SEM). Figure 4B shows that attentional improvement in neuronal d' (attended-unattended) for each hemisphere-day is highly correlated with behavioural improvement ($R=0.69$, $p<10^{-12}$). This strong correlation suggests that our d' metric captures the important aspects of the improvements in population sensitivity that lead to improvements in behaviour.

Each of the physiological changes we observed in rate, Fano factor, and correlation could have contributed to the improvement in population sensitivity. We next compared how much attentional modulation of each factor alone and the three factors together contributed to the actual improvements in d' that we calculated. To isolate the contribution of each factor, we simulated the responses of populations of neurons using the same mean rates, noise correlations, Fano factors, and number of neurons as the groups of neurons we recorded (for simulation methods, see Supplementary information) and compared the calculated d' for each simulation to the real data in the unattended condition. In this example data set, attentional modulation of all three factors together (top right of Figure 4C) accounts for 95% of the attentional improvement we observed in the real data. We then calculated the contribution of each factor separately by simulating attentional modulation of the factor of interest and using the values observed in the unattended condition for the other two factors (second through fourth rows of the right column Figure 4C). Correlation alone accounted for 79% of the attentional improvement, rate accounted for 9%, and modulation of independent variability accounted for 4%.

The example in Figure 4C is typical of the 82 data sets. On average, attentional modulation of the three factors together accounted for 92% of the attentional improvement we observed in the actual populations (black bar in Figure 4D). Importantly, this result means that population sensitivity is well modeled by accounting only for rate, independent variability, and pairwise noise correlation, and that any other factors (including any higher order correlations) account for no more than 8% of the observed improvement in population sensitivity. Consistent with this, population responses in the retina are well described by the responses of individual neurons and pairwise correlations^{35, 36}. Overall, modulation of noise correlation was by far the most important factor in explaining the improvement in population sensitivity. Attentional modulation of noise correlation accounted for 81% of the observed improvement, rate accounted for 10%, and independent variability accounted for only 0.3% (which was not significantly different than 0.0; t-test, $p=0.82$).

Unsurprisingly, we found that both the observed raw population d' and the improvement in d' due to attention depended on the number of neurons we recorded. Because there is no a priori way of knowing how many neurons are involved in the task, we examined the dependence of these measures on population size by sampling, with replacement, the firing rates, Fano factors, and correlations of all of the neurons we recorded over all recording sessions (see Supplementary information and^{11, 24} for methods). Figure 4E shows population d' for simulations in which attention modulated either all three factors, one factor individually, or none of the factors. In all cases, d' increases with population size. Because noise correlations are on average positive for both attentional conditions, d' asymptotes for large populations⁹⁻¹¹. Modulation of noise correlation accounts for most of the attentional improvement in sensitivity across nearly all population sizes (difference between the colored lines and the grey “unattended” line in Figure 4E). For very small populations, however, this is necessarily not true (inset, Figure 4E). If performance depends on a single neuron, there can be no correlation, and the small attentional improvement depends almost entirely on modulation of firing rate. In our simulations, noise correlations become dominant for populations of more than five neurons. If anything, this estimate may be high because we recorded from many neurons with stimulus preferences that were not well matched to the

stimuli we presented, resulting in low firing rate responses and low noise correlations (Figure 2C). If we had recorded from neurons better matched to the stimuli, correlations would likely have been higher (right side of Figure 2C), shifting the point at which correlation becomes most important to population sizes even lower than five neurons. Many more than five neurons are thought to be involved in virtually every task, so changes in correlation likely dominate attentional improvement in nearly all situations.

Discussion

Why do changes in shared variability have a bigger impact on population sensitivity than changes in the signal-to-noise of single neurons? One answer is that the changes in correlated variability we observed were larger than the changes in firing rate or Fano factor. However, we re-ran the simulations in Figure 4E assuming that the three factors all had the same modulation index as the changes in rate (see Supplementary information), and correlation still dominated for population sizes greater than 30 neurons. Instead, the explanation lies in the fact that no matter how noisy individual neurons are, independent variability can be averaged out if the population of neurons is large enough. Correlated variability, however, can never be averaged out by simply adding neurons to the population.

Noise correlation can either improve or reduce population sensitivity, depending on the algorithm by which neural responses are read out^{8, 9}, and our simulations could in principle have revealed that the observed correlation decreases acted to reduce population sensitivity. However, theoretical studies show that decreased correlation improves discrimination if the difference between the responses to the stimuli to be discriminated (the original and changed stimuli in our task) are of the same sign for most neurons^{8, 9}, which turned out to be the case in our data set.

Most of the neurons we recorded (92%) responded more strongly to the changed than the previous stimulus, presumably reflecting adaptation to the series of identical stimuli preceding the change. Therefore, the optimal quantity to be read out is similar to a (positively) weighted mean of the responses of the population, and indeed, the axis of discrimination we determined using the procedure depicted in Figure 4A was close to the weighted population mean. The attention-related decrease in correlation therefore improved the sensitivity of the population by reducing the amount of shared variability that could not be removed by averaging. In contrast, a recent study showed no effect of attention on noise correlations in a situation in which correlations were shown to have no effect on the sensitivity of the population³⁷. There are further situations (such as those in which the optimal readout algorithm is more similar to a subtraction of two populations of cells) in which an increase in correlation would improve the sensitivity of the population^{8, 9}. Whether attention would increase correlations in such tasks remains to be determined.

It is likely, of course, that the brain uses a different algorithm for extracting stimulus information from the responses of many neurons than the very simple decoding scheme we used in the simulations in Figure 4. However, the observation that attentional modulation of noise correlation explains most of the attentional improvement in population sensitivity is likely true for any sensible decoding algorithm. First, the difference in the amount of

attentional improvement explained by pairwise correlations is very large compared to the amount explained by the changes in the responses of single neurons, suggesting that noise correlations will dominate using any decoding algorithm. Furthermore, correlation was by far the most important factor using any of several linear discriminators we tried, including the single axis projection described here, Fisher discriminants, and support vector machines (data not shown). Higher order decoders that explicitly read out interneuronal correlations^{38_41} will be even more affected by attentional modulation of correlation than linear discriminators. Finally, any sort of decoding algorithm that incorporates a mean (or weighted mean) of the responses of many neurons will be greatly affected by noise correlations^{10_12}.

Mathematically, correlation is invariant to the mean response (the correlation coefficient is the ratio of the covariance to the square root of the product of the individual variances, so both the numerator and denominator are proportional to the product of the means), so underlying noise correlations cannot be changed by a simple scaling of neural responses (i.e. a gain change). Instead, noise correlations in cortex are thought to arise primarily from common, noisy inputs^{10, 22, 23, 29}. The fact that attention primarily decreases correlations provides clues about the mechanisms by which attention affects populations of sensory neurons. A decrease in correlation combined with an increase in firing rates is consistent with a decrease in the strength of an effectively inhibitory input that is common across the population. One possibility is that attention results in a decrease in the weights or activity of inputs that cause divisive normalization, a mechanism that normalizes responses to many stimuli within a receptive field and has recently been proposed to underlie attention^{34, 42, 43}. In fact, we found a correlation between the mean attentional modulation of the firing rates of a pair of neurons and modulation of their noise correlation ($R=-0.32$, $p<10^{-4}$) and also between the average rate and correlation changes within a hemisphere-day ($R=-0.61$, $p<10^{-9}$; see Supplementary Figure 2), which is consistent with the idea that the two attentional changes may be mediated by the same mechanism.

Attention improves perception of the attended location or feature, so studying the effects of attention on populations of sensory neurons reveals the aspects of the population code that are most important for accurately encoding information about a behaviourally relevant stimulus. We have shown here that attention improves population sensitivity primarily by changing noise correlations, and even the small pairwise correlations we observed have a dramatic effect on the sensitivity of the population. Therefore, understanding the interactions between pairs of neurons is critically important for understanding population coding (see also^{8_11, 38_41, 44}).

Rather than examining mean responses over many trials, the brain makes decisions based on the responses of many neurons over a short period. Our results show that studies of average responses of single neurons miss interactions between neurons that have critical effects on behaviour. Together, these results suggest that the future of studying population coding will rely on multi-electrode or imaging technologies that allow glimpses of population coding on the timescale of a single behavioural decision.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Douglas Ruff for spike sorting and animal training assistance, Mark Churchland for code and advice regarding mean-matching and Fano factor analysis, and Mark Histed for many helpful discussions. They, Amy Ni, and Alexandra Smolyanskaya provided comments on an earlier version of the manuscript. This work was supported by NIH grant R01EY005911.

References

1. Yantis S, Serences JT. Cortical mechanisms of space-based and object-based attentional control. *Curr Opin Neurobiol.* 2003; 13:187–193. [PubMed: 12744972]
2. Assad JA. Neural coding of behavioural relevance in parietal cortex. *Curr Opin Neurobiol.* 2003; 13:194–197. [PubMed: 12744973]
3. Reynolds JH, Chelazzi L. Attentional modulation of visual processing. *Annu Rev Neurosci.* 2004; 27:611–647. [PubMed: 15217345]
4. Maunsell JH, Treue S. Feature-based attention in visual cortex. *Trends Neurosci.* 2006; 29:317–322. [PubMed: 16697058]
5. Tolhurst DJ, Movshon JA, Dean AF. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* 1983; 23:775–785. [PubMed: 6623937]
6. McAdams CJ, Maunsell JH. Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron.* 1999; 23:765–773. [PubMed: 10482242]
7. Mitchell JF, Sundberg KA, Reynolds JH. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron.* 2007; 55:131–141. [PubMed: 17610822]
8. Abbott LF, Dayan P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* 1999; 11:91–101. [PubMed: 9950724]
9. Averbach BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nat Rev Neurosci.* 2006; 7:358–366. [PubMed: 16760916]
10. Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature.* 1994; 370:140–143. [PubMed: 8022482]
11. Shadlen MN, Britten KH, Newsome WT, Movshon JA. A computational analysis of the relationship between neuronal and behavioural responses to visual motion. *J Neurosci.* 1996; 16:1486–1510. [PubMed: 8778300]
12. Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci.* 1998; 18:3870–3896. [PubMed: 9570816]
13. Desimone R, Schein SJ. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol.* 1987; 57:835–868. [PubMed: 3559704]
14. Gattass R, Sousa AP, Gross CG. Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci.* 1988; 8:1831–1845. [PubMed: 3385477]
15. Motter BC. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol.* 1993; 70:909–919. [PubMed: 8229178]
16. McAdams CJ, Maunsell JH. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci.* 1999; 19:431–441. [PubMed: 9870971]
17. Williford T, Maunsell JH. Effects of spatial attention on contrast response functions in macaque area V4. *J Neurophysiol.* 2006; 96:40–54. [PubMed: 16772516]
18. Treue S, Martinez Trujillo JC. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature.* 1999; 399:575–579. [PubMed: 10376597]

19. Churchland, MM., et al. Computational and systems neuroscience. *Frontiers in Systems Neuroscience*; Salt Lake City, UT: 2009. Stimulus onset quenches neural variability: a widespread cortical phenomenon.
20. Churchland MM, Yu BM, Sahani M, Shenoy KV. Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Curr Opin Neurobiol.* 2007; 17:609–618. [PubMed: 18093826]
21. Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science.* 2001; 291:1560–1563. [PubMed: 11222864]
22. Kohn A, Smith MA. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J Neurosci.* 2005; 25:3661–3673. [PubMed: 15814797]
23. Bair W, Zohary E, Newsome WT. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci.* 2001; 21:1676–1697. [PubMed: 11222658]
24. Cohen MR, Newsome WT. Context-dependent changes in functional circuitry in visual area MT. *Neuron.* 2008; 60:162–173. [PubMed: 18940596]
25. Thut G, Nietzel A, Brandt SA, Pascual-Leone A. Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci.* 2006; 26:9494–9502. [PubMed: 16971533]
26. Foxe JJ, Simpson GV, Ahlfors SP. Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport.* 1998; 9:3929–3933. [PubMed: 9875731]
27. Babiloni C, et al. Sub-second “temporal attention” modulates alpha rhythms. A high-resolution EEG study. *Brain Res Cogn Brain Res.* 2004; 19:259–268. [PubMed: 15062863]
28. Bastiaansen MC, Bocker KB, Brunia CH, de Munck JC, Spekreijse H. Event-related desynchronization during anticipatory attention for an upcoming stimulus: a comparative EEG/MEG study. *Clin Neurophysiol.* 2001; 112:393–403. [PubMed: 11165546]
29. Smith MA, Kohn A. Spatial and temporal scales of neuronal correlation in primary visual cortex. *J Neurosci.* 2008; 28:12591–12603. [PubMed: 19036953]
30. de la Rocha J, Doiron B, Shea-Brown E, Josic K, Reyes A. Correlation between neural spike trains increases with firing rate. *Nature.* 2007; 448:802–806. [PubMed: 17700699]
31. Reynolds JH, Pasternak T, Desimone R. Attention increases sensitivity of V4 neurons. *Neuron.* 2000; 26:703–714. [PubMed: 10896165]
32. Roelfsema PR, Spekreijse H. The representation of erroneously perceived stimuli in the primary visual cortex. *Neuron.* 2001; 31:853–863. [PubMed: 11567622]
33. Sundberg KA, Mitchell JF, Reynolds JH. Spatial attention modulates center-surround interactions in macaque visual area v4. *Neuron.* 2009; 61:952–963. [PubMed: 19324003]
34. Lee J, Maunsell JH. A normalization model of attentional modulation of single unit responses. *PLoS ONE.* 2009; 4:e4651. [PubMed: 19247494]
35. Shlens J, et al. The structure of multi-neuron firing patterns in primate retina. *J Neurosci.* 2006; 26:8254–8266. [PubMed: 16899720]
36. Schneidman E, Berry MJ 2nd, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature.* 2006; 440:1007–1012. [PubMed: 16625187]
37. Poort J, Roelfsema PR. Noise correlations have little influence on the coding of selective attention in area V1. *Cereb Cortex.* 2009; 19:543–553. [PubMed: 18552357]
38. Series P, Latham PE, Pouget A. Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nat Neurosci.* 2004; 7:1129–1135. [PubMed: 15452579]
39. Beck JM, et al. Probabilistic population codes for Bayesian decision making. *Neuron.* 2008; 60:1142–1152. [PubMed: 19109917]
40. Pouget A, DeAngelis GC. Paying attention to correlated neural activity. *Nat Neurosci.* 2008; 11:1371–1372. [PubMed: 19023342]
41. Pillow JW, et al. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature.* 2008; 454:995–999. [PubMed: 18650810]

42. Reynolds JH, Heeger DJ. The normalization model of attention. *Neuron*. 2009; 61:168–185. [PubMed: 19186161]
43. Boynton GM. A framework for describing the effects of attention on visual responses. *Vision Res*. 2009; 49:1129–1143. [PubMed: 19038281]
44. Kohn A, Zandvakili A, Smith MA. Correlations and brain states: from electrophysiology to functional imaging. *Curr Opin Neurobiol*. 2009

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

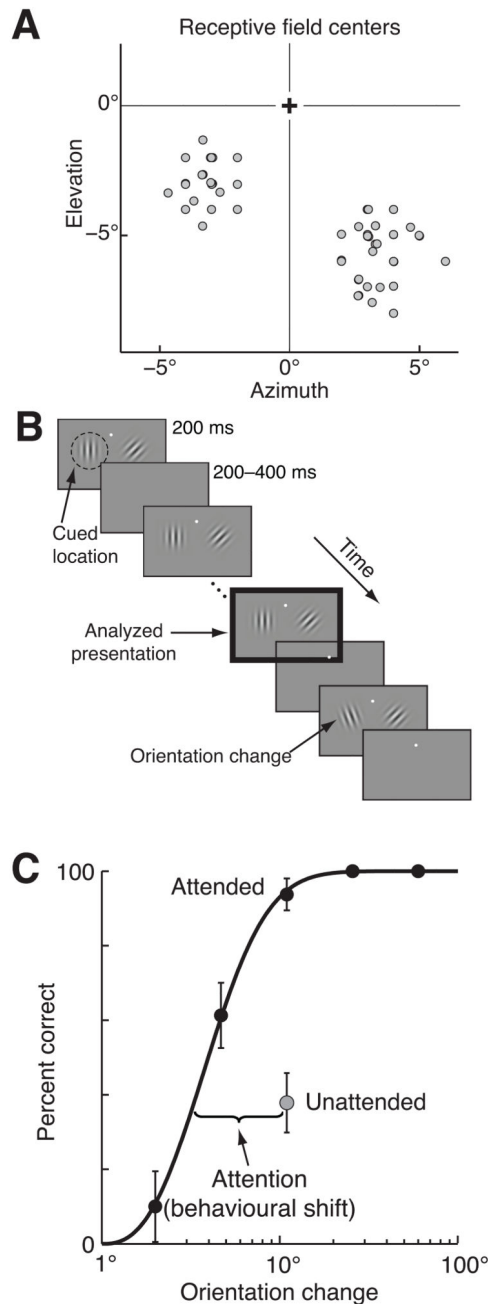


Figure 1.

Methods and behaviour. **A.** Center of visual receptive fields for the multiunit signals from one monkey. **B.** Orientation change detection task. Two Gabor stimuli synchronously flashed on for 200 ms and off for a randomized 200–400 ms period. At an unsignaled and randomized time, the orientation of one of the stimuli changed, and the monkey was rewarded for making a saccade to the stimulus that changed. Attention was cued in blocks, and the cue was valid on 80% of trials, meaning that on an “attend-left” block of trials (depicted here), 80% of orientation changes were to the left stimulus. The monkey was rewarded for correctly detecting any change, even on the unattended side. Unless otherwise

stated, all analyses were performed on responses to the stimulus before the orientation change (black outlined panel). **C.** Psychometric performance from a typical example experiment. Proportion correct as a function of orientation change in degrees for trials in which the change occurred at the attended (black points) or unattended (grey point) location. Unattended changes occurred only at the middle difficulty level (11°). Attentional improvement in behaviour was quantified as the lateral shift between the percent correct on unattended trials and the fitted psychometric curve for attended trials.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

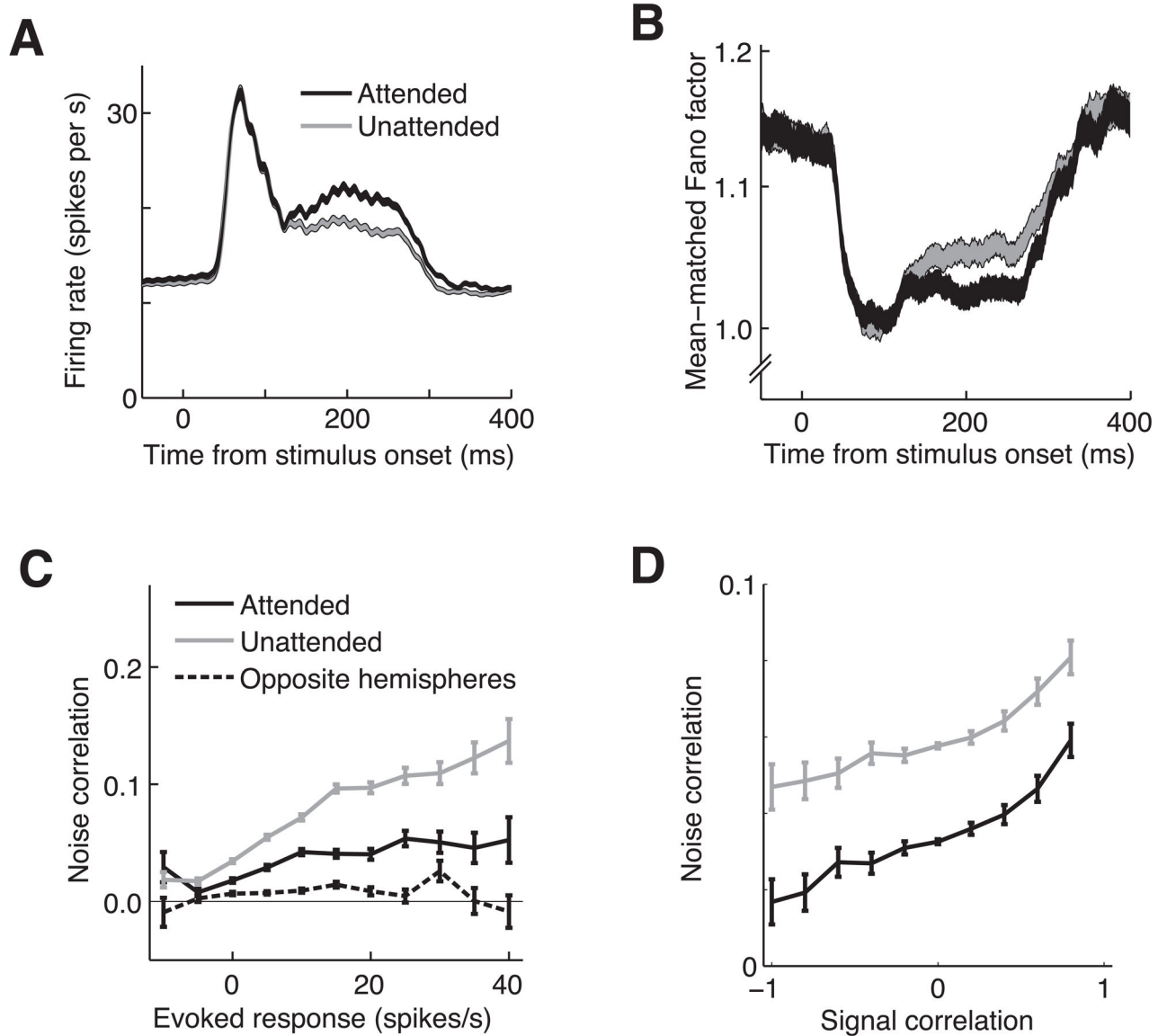


Figure 2.

Attentional modulation of firing rate, Fano factor, and noise correlation. **A.** Attention increases firing rates. Peri-stimulus time histogram of firing rates for all 3,498 single neurons and multiunit clusters on trials when the stimulus in the same hemifield as the neuron's receptive field was attended (black line) or unattended (grey line). Line width represents the SEM. **B.** Attention decreases mean-matched Fano factor. Plotting conventions are as in A. **C.** Attention decreases noise correlation. Spike count noise correlation (for responses over the period from 60 to 260 ms following stimulus onset) is plotted as a function of the mean stimulus modulation for the pair of neurons (firing rate during the stimulus – firing rate during the interstimulus blank period). For pairs of neurons in the same hemisphere, correlation was lower when the stimulus in the neurons' receptive field was attended (black line) than when it was unattended (grey line). Pairs of neurons in opposite hemispheres (dashed lines) had correlations that were close to zero. Error bars represent

SEM. **D.** Raw noise correlation, but not attentional modulation, signal correlation depends on signal correlation. Mean noise correlation is plotted as a function of signal correlation, which can be thought of as the similarity in spatial and feature tuning of the two neurons (see Methods). As has been previously reported, noise correlation increases with signal correlation. However, the difference in correlation between the attended (black line) and unattended (grey line) conditions did not depend on signal correlation. Error bars represent SEM.

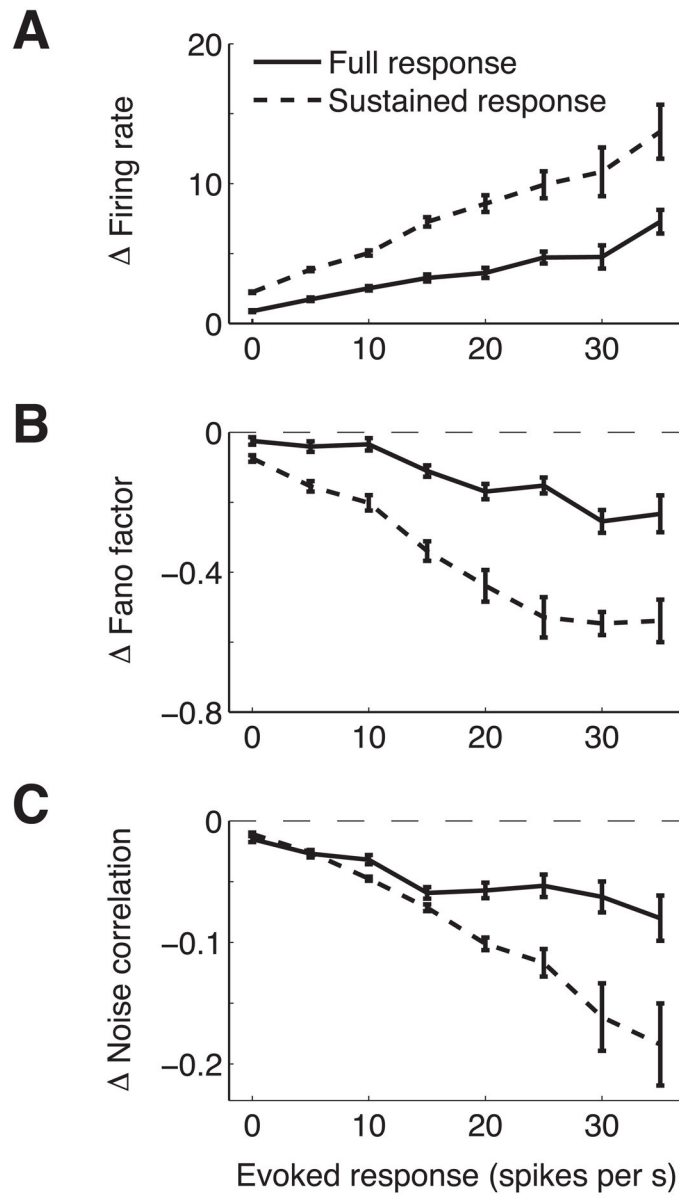


Figure 3.

Attention has the biggest effects on the most responsive neurons. **A.** Difference in mean firing rate between trials when the stimulus in the neuron's receptive field was attended and unattended as a function of stimulus modulation (rate during stimulus period – interstimulus period). Error bars represent SEM. **B.** Same, for Fano factor. **C.** Same, for noise correlation for pairs of neurons in the same hemisphere.

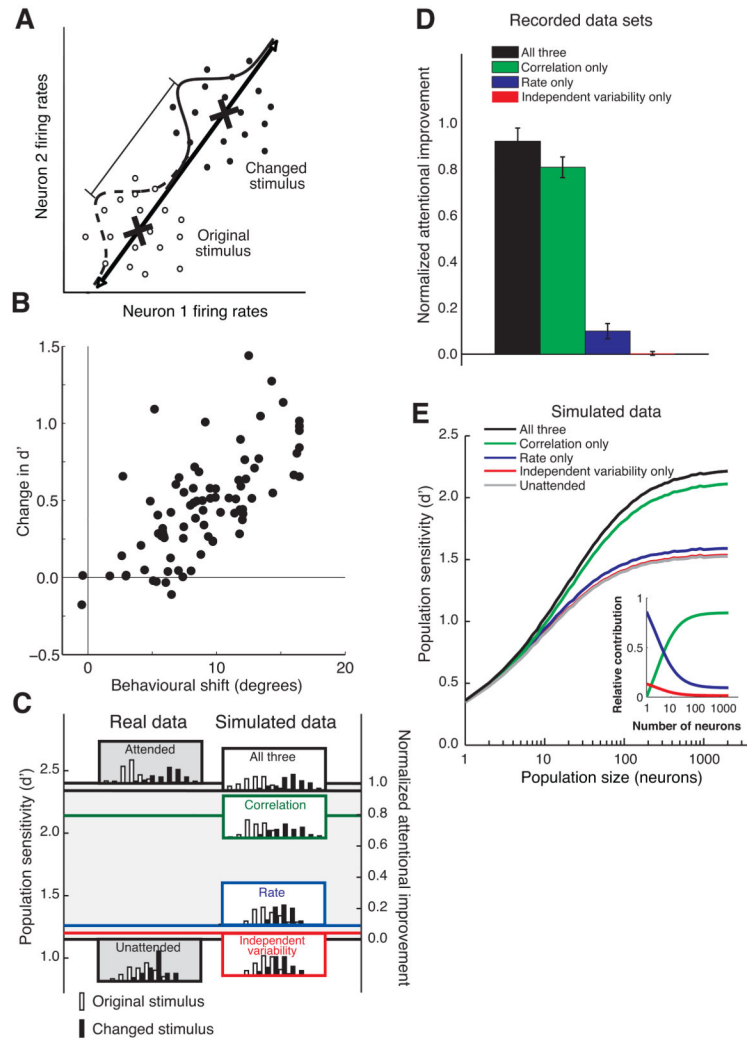


Figure 4. Modulation of noise correlation accounts for the majority of the attentional improvement in population sensitivity. **A.** Procedure for calculating the sensitivity of the population. For each trial and attentional condition, the firing rate response of the n neurons recorded simultaneously in a given hemisphere to the stimulus immediately before the orientation change (open circles) and the changed stimulus (filled circles) is plotted as a point in an n -dimensional space (a fictional two-neuron example is plotted here). The axis of discrimination (black line) is the line connecting center of mass of the n -dimensional cloud of points for each time period (X's). Each point is projected onto the axis, leaving a one-dimensional distribution of projected values for each time period (dashed and solid curves). The sensitivity of the population to the change in the stimulus is quantified as the discriminability of the two distributions in units of d' (the difference between changed mean and original mean divided by the standard deviation). **B.** Population d' and behavioral improvement are highly correlated. For each hemisphere-day, population d' is plotted as a function of the behavioral improvement (quantified as the lateral shift between performance at the unattended location and the fitted psychometric curve for the attended condition). **C.**

Procedure for calculating the amount of the observed attentional improvement explained by each factor for a representative example data set. Histograms of projections onto the axis defined in A are plotted for the real data (left column, for attended and unattended trials), and for simulations (right column). We defined the observed attentional improvement as the difference between the d' 's for the two attentional conditions ($d'=2.40$ for the attended condition and 1.15 for the unattended condition, giving an improvement of 1.25 in this example). The left axis represents d' and the right axis represents normalized proportion of attentional improvement (by definition 1.0 for the attended condition and 0.0 for the unattended condition). To isolate the contribution of each factor (or group of factors), we simulated responses of an identically sized population of neurons with the same mean firing rate, Fano factor, and noise correlation as each of the neurons in our data set in which the statistics of the labeled factor/s matched the data for the attended condition and the other factors matched the data for the unattended condition (right column of distributions). We calculated the fraction of the observed attentional improvement explained by each factor/s by comparing the simulated d' to the d' for the real unattended data. In this data set, modulation of independent variability (at the level predicted if changes in Fano factor were due solely to changes in independent variability) accounted for 4% of the observed attentional improvement, rate accounted for 9%, correlation accounted for 79%, and the three together accounted for 95%. **D.** Average proportion of actual attentional improvement for all 82 data sets. Each day of data contributed two data sets (one for each hemisphere). Error bars represent SEM. All proportions are statistically different from zero (t-test, $p<0.01$) except the independent variability-only simulation ($p=0.82$). **E.** Population sensitivity as a function of the number of neurons involved in the task. Population d' was calculated using the method described in A and B except that data in both the attended and the unattended conditions were simulated. For each population size, we sampled, with replacement, from the entire population of neurons from all data sets combined. Each simulation was run 100 times for 10,000 trials on each run. The inset plots the relative contribution of each factor (which is the ratio of the improvement in d' for that factor alone to the improvement in d' for all three factors) as a function of population size. Correlation is the most important factor for population sizes greater than 5 (crossing of the green and blue lines).