Article

Neuron

# Attentional Changes in Either Criterion or Sensitivity Are Associated with Robust Modulations in Lateral Prefrontal Cortex

## Highlights

- Two behavioral components of visuospatial attention were isolated
- Visual responses in lateral prefrontal cortex were robustly modulated by either component
- Modulations in prefrontal cortex and visual cortex were qualitatively different
- Separate brain areas are associated with distinct mechanisms of attention

## Authors

Thomas Zhihao Luo,
John H.R. Maunsell

## Correspondence

thomas.zhihao.luo@gmail.com (T.Z.L.),
maunsell@uchicago.edu (J.H.R.M.)

## In Brief

Luo and Maunsell show that the modulations in prefrontal cortex correspond to multiple components of attention and differ from modulations in visual cortex, indicating that different brain structures underlie distinct attentional mechanisms and that attention is not a unitary process.

Luo & Maunsell, 2018, Neuron *97*, 1382–1393
March 21, 2018 
https://doi.org/10.1016/j.neuron.2018.02.007

CellPress

CellPress

Neuron
Article

# Attentional Changes in Either Criterion or Sensitivity Are Associated with Robust Modulations in Lateral Prefrontal Cortex

Thomas Zhihao Luo[1,2,3,*] and John H.R. Maunsell[1,*]
[1]Department of Neurobiology, the University of Chicago, Chicago, IL 60637, USA
[2]Present address: Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, USA
[3]Lead Contact
*Correspondence: thomas.zhihao.luo@gmail.com (T.Z.L.), maunsell@uchicago.edu (J.H.R.M.)
https://doi.org/10.1016/j.neuron.2018.02.007

## SUMMARY

**Visual attention is associated with neuronal changes across the brain, and these widespread signals are generally assumed to underlie a unitary mechanism of attention. However, using signal detection theory, attention-related effects on performance can be partitioned into changes in either the subject's criterion or sensitivity. Neuronal modulations associated with only sensitivity changes were previously observed in visual cortex, raising questions about which structures mediate attention-related changes in criterion and whether individual neurons are involved in multiple components of attention. Here, we recorded from monkey lateral prefrontal cortex (LPFC) and found that, in contrast to visual cortex, neurons in LPFC changed their firing rates, pairwise correlation, and Fano factor when subjects changed either their criterion or their sensitivity. These results indicate that attention-related neuronal modulations in separate brain regions are not a monolithic signal and instead can be linked to distinct behavioral changes.**

## INTRODUCTION

When attention is directed to a location in visual space, the activity of visually responsive neurons changes throughout the brain (Krauzlis et al., 2013; Maunsell, 2015; Moore and Zirnsak, 2017). When attention is directed within a neuron's receptive field, its firing rate typically increases relative to when attention is elsewhere. Changes in firing rates associated with visuospatial attention are observed in the visual, parietal, and prefrontal regions of the cerebral cortex (Wurtz and Mohler, 1976; Lynch et al., 1977; Moran and Desimone, 1985), as well as in the superior colliculus and thalamus (Ignashchenkova et al., 2004; McAlonan et al., 2008; Briggs et al., 2013). Except for the retina, every brain region examined that contains neurons with spatially selective visual responses has been shown to contain signals related to visuospatial attention. While these modulations have been characterized by many studies since first observed in the superior colliculus (Goldberg and Wurtz, 1972), it remains unclear how they give rise to the behavioral effects associated with visuospatial attention. In particular, it is unknown whether distinct behavioral changes related to visuospatial attention depend on separate or on the same brain areas.

When attention is directed to a visual location, an observer detects a higher proportion of targets at that location and responds to targets there with shorter delays (Posner et al., 1980; Bashinski and Bacharach, 1980). When the target is barely visible, attention can reliably make the difference between a successful detection and failure. In single-neuron recording studies, when an animal shows an increase in target detection rate (i.e., hit rate) without changes in the stimulus or eye movements, concomitant neuronal changes are considered to be correlates of visuospatial attention.

However, the attention-related improvement in hit rate can be partitioned into two independent components using signal detection theory: either changes in the subject's criterion ($c$) or changes in the subject's sensitivity ($d'$) (Figure 1A) (Bashinski and Bacharach, 1980; Müller and Findlay, 1987; Downing, 1988; Hawkins et al., 1990; Müller and Humphreys, 1991; Wyart et al., 2012; Luo and Maunsell, 2015; Sridharan et al., 2017; Arcizet et al., 2017). A subject's criterion corresponds to how readily the subject reports a target or withholds from such a report. Adopting a more liberal criterion (a decrease in $c$) at a visual location consists in responding more frequently at that location, resulting in more targets being detected there even when the stimulus is unchanged. A subject's sensitivity corresponds to the ability to discriminate between a target and a nontarget. Enhancing sensitivity (increasing $d'$) at a location also results in more targets detected at that location. The distinction between criterion and sensitivity is crucial because any improvement in an observer's hit rate can be equivalently brought about by a decrease in $c$ or an increase in $d'$ (Figure 1B). These two changes are differentiated by the subject's false alarm rate. An improvement in hit rate brought about by a decrease in $c$ is associated with a higher false alarm rate, while the same increase in hit rate brought about by an increase in $d'$ is associated with a lower false alarm rate.

Because single-neuron studies typically have operationalized visuospatial attention using only hit rates, it is unclear how

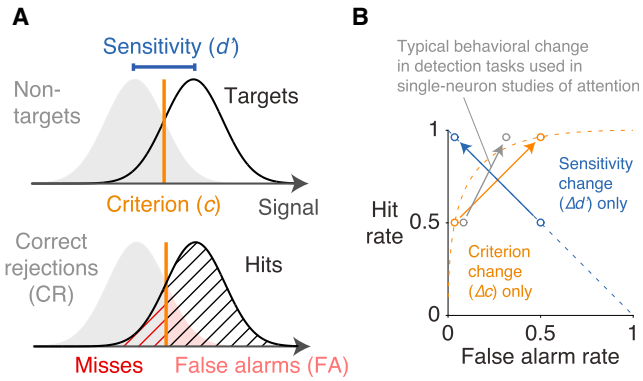


**Figure 1. Visuospatial Attention Can Be Partitioned Using Signal Detection Theory**

(A) In the signal detection model, each stimulus evokes a noisy internal signal. If the signal is stronger than the criterion (*c*), the stimulus is reported as a target. The distributions of internal signals evoked by the target and by the nontarget overlap, and the separation between these two distributions is indexed as sensitivity (*d'*). The subject's response to each stimulus is categorized as a hit (H), miss (M), false alarm (FA), or correct rejection (CR), and the frequency of these responses are used to calculate *c* and *d'*.

(B) An increase in hit rate, which is used to operationalize attention in many single-neuron studies, typically consists of a combination of changes in criterion and sensitivity (Luo and Maunsell, 2015). A typical behavioral change of this sort is shown (in gray), along with an isolated change in criterion (in orange) and an isolated change in sensitivity (in blue). The dashed orange line represents an isopleth along which *d'* is constant and only *c* varies, and the dashed blue line represents the isopleth along which *c* is constant and only *d'* varies. Any change in hit rate can be brought about by a change in criterion alone, a change in sensitivity alone, or both.

signals across the brain relate to these two components of attention. We previously designed a task for monkeys that isolated these two behavioral changes and found that neuronal modulations in visual cortical area V4 are associated with changes in sensitivity, but not with changes in criterion (Luo and Maunsell, 2015). This previous result indicates that at least in this task, changes in behavioral criterion depend on brain regions other than V4, and therefore suggests that separate brain regions contribute in distinct ways to visuospatial attention. Consistent with this interpretation, a recent study suggests that the superior colliculus contributes to visuospatial attention primarily through changes in the subject's criterion rather than sensitivity (Sridharan et al., 2017).

Yet it remains unclear what neuronal signals are associated with attention-related changes in criterion, and whether criterion-related neuronal changes are qualitatively similar to the signals typically correlated with attention. To address these questions, we recorded from neurons in lateral prefrontal cortex (LPFC) while monkeys changed either their criterion or sensitivity at specific visual locations. Neuronal changes in LPFC have been correlated with spatially selective changes in response readiness (Boch and Goldberg, 1989; Boussaoud and Wise, 1993) and in target detection (Lennert and Martinez-Trujillo, 2011). However, because previous studies recording from LPFC have not controlled the subject's criterion and sensitivity, it is unknown to which component of attention the modulations in LPFC correspond.

We found that visual responses in LPFC were robustly modulated when the subject changed either criterion or sensitivity. Because the modulations in LPFC differ qualitatively from those in visual cortical area V4, which are correlated with changes only in sensitivity, the results demonstrate that attention-related modulations in different brain structures are not a unitary signal and instead relate to distinct behavioral components of attention.

## RESULTS

### Isolation of Behavioral Changes in Either Criterion or Sensitivity

We trained two monkeys on a change detection task that isolated spatially selective changes in either the subject's criterion or its sensitivity (Figure 2A) (Luo and Maunsell, 2015). In each trial, two stimuli ("samples") appeared concurrently for 400 ms. After a delay of 150–250 ms, a single stimulus ("test") appeared at one of the two sample locations selected at random. The monkey had to saccade to the test if it differed in orientation from the sample at the same location. The probability of an orientation change between the samples and the test was 0.5. If no change occurred, the monkey had to wait and saccade to a second test stimulus, which always differed from the sample. The response to the first test in each trial was categorized as a hit (H), miss (M), false alarm (FA), or correct rejection (CR), and these responses were used to compute *c* and *d'*.

To control the subject's criterion and sensitivity, we titrated the reward given for a H or a CR separately at each stimulus location. The subject's criterion was primarily controlled by the ratio of the reward for a H to the reward for a CR ("H:CR reward ratio," Figures 2B and 2C). A larger H:CR reward ratio at a location encouraged a more lenient criterion (i.e., lower or more negative *c*) at that location, while a smaller H:CR reward ratio encouraged a more stringent criterion (i.e., higher or more positive *c*). The subject's sensitivity was primarily controlled by the average reward across H and CR. A larger average reward at a location results in a higher sensitivity at that location. These reward contingencies were varied between two task conditions of each daily session to isolate a change in either the subject's criterion (in a "$\Delta c$ isolation session") or its sensitivity (in a "$\Delta d'$ isolation session") at each stimulus location.

We isolated behavioral changes in 48 sessions: 12 sessions of criterion change and 12 sessions of sensitivity change performed by each of two monkeys (Figures 2C–2F). Three other sessions in which we failed to adequately isolate a behavioral change were excluded. Behavioral changes were spatially specific, as demonstrated by the counterphase change in behavior between the two stimulus locations (Figure S1A), and therefore not due to global changes, such as arousal. Similar changes in hit rate were obtained with sensitivity changes and criterion changes (Figure 2D). Behavioral control was prioritized at the stimulus location contralateral to the electrode array at the ipsilateral location (STAR Methods). Among the 48 sessions with satisfactory isolation at the contralateral location, 33 sessions also achieved isolation at the ipsilateral location (Figure S1B).

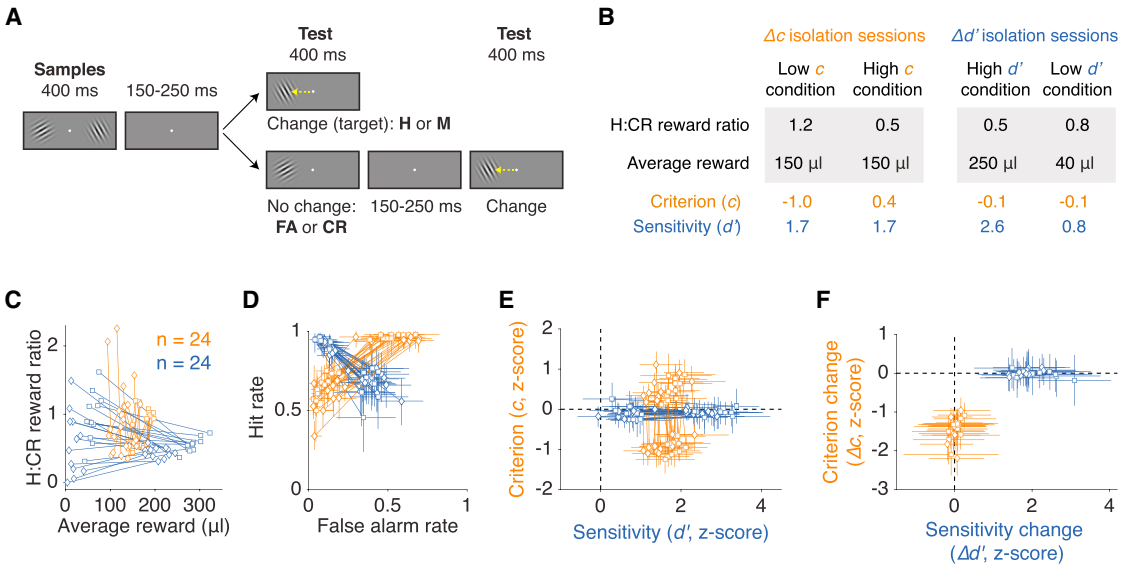


**Figure 2. Isolation of Attentional Changes in Criterion and Sensitivity**

(A) Monkeys had to detect an orientation change that occurred on either the first or second test stimulus and report it with a saccade. If no change occurred on the first test, monkeys had to wait for a change that would always occur on the second test. In a trial in which a change occurred on the first test, the outcome was categorized as either a hit (H) or a miss (M), and when no change occurred on the first test, the outcome was either a false alarm (FA) or a correct rejection (CR).

(B) Average reward parameters and behavioral indices across sessions. Each session consists of two task conditions, and isolation was achieved by varying reward sizes between the two conditions. Criterion was controlled primarily using the ratio of reward for a H to a reward for a CR ("H:CR reward ratio"), and sensitivity was controlled primarily using the average reward across H and CR. The task condition (*high d'*, *low d'*, *low c*, or *high c*) indicates the animal's performance at the contralateral location.

(C) Reward contingencies in each daily session. Each marker refers to one of the two task conditions of each session, and the pair of markers representing the same session are connected by a line. The data are from 24 $\Delta c$ isolation sessions (shown in orange) and 24 $\Delta d'$ isolation sessions (in blue). Diamonds and squares represent monkeys C and Y, respectively.

(D) Hit rates and false alarm rates of each task condition of each daily session.

(E) Criterion and sensitivity of each task condition of each daily session.

(F) Criterion change and sensitivity change in each session. Crosshairs in (D)–(F) represent 95% confidence intervals.

See also Figure S1.

## Neurons Were Recorded from Lateral Prefrontal Cortex and Tracked across Days

In the right hemisphere of each monkey, we implanted a pair of 6 × 8 Utah microelectrode arrays anterior to the arcuate sulcus, either above or below the principal sulcus. The arrays overlapped with Walker's areas 8A, 45, and 46 (Walker, 1940) and, according to the map by Petrides and Pandya (1999), the areas 8Ad, 8Av, 9/46v, 45B, and 45A (Figure S2). We refer to these areas as lateral prefrontal cortex (LPFC).

We recorded from neuronal units during the 48 behavioral isolations described above. In each daily session, two Gabor stimuli were placed in opposite hemifields, and their positions and orientations varied across days. On a typical day, we recorded from 94 units (11 fully isolated single units and 83 partially isolated units or multiunit clusters).

Because it is highly likely that the Utah arrays recorded from the same neurons across days (Dickey et al., 2009), we implemented the classification algorithm by Fraser and Schwartz (2012) to determine whether units in separate sessions represent the same neurons (STAR Methods). The classifier was calibrated to produce similar error rates for two types of errors: a decoy error in which different neurons were misclassified as the same neuron, and a drop error in which the same neuron was misclassified as two different neurons. The chosen classifier had a median decoy error rate of 0.0125 and drop error rate of 0.0121(Figures S3A and S3B). Sixty-two percent of units were recorded in more than one session, and the same unit could be recorded in sessions more than 30 days apart (Figures S3C and S3D). Among 4,843 units encountered, 1,088 were classified as unique.

## Spatial Selectivity Was Indexed Using a Memory-Guided Saccade Task

Neurons in LPFC typically have extensive receptive fields that overlap with both visual hemifields (Funahashi et al., 1990). To study neuronal activity related to spatial attention, we measured the degree to which each neuron differentiated the two stimulus locations. At the start of each day, the animal performed a memory-guided saccade task. In this task, a target appeared for 400 ms in either the left or right hemifield, at the centers of the two Gabors subsequently displayed in the attention task. After a delay of 750 to 1,000 ms, the fixation point disappeared, and the animal saccaded to the remembered target location for a reward.

We computed a spatial selectivity index (SSI) for each neuron based on its activity after the target onset ($SSI_{visual}$), during the

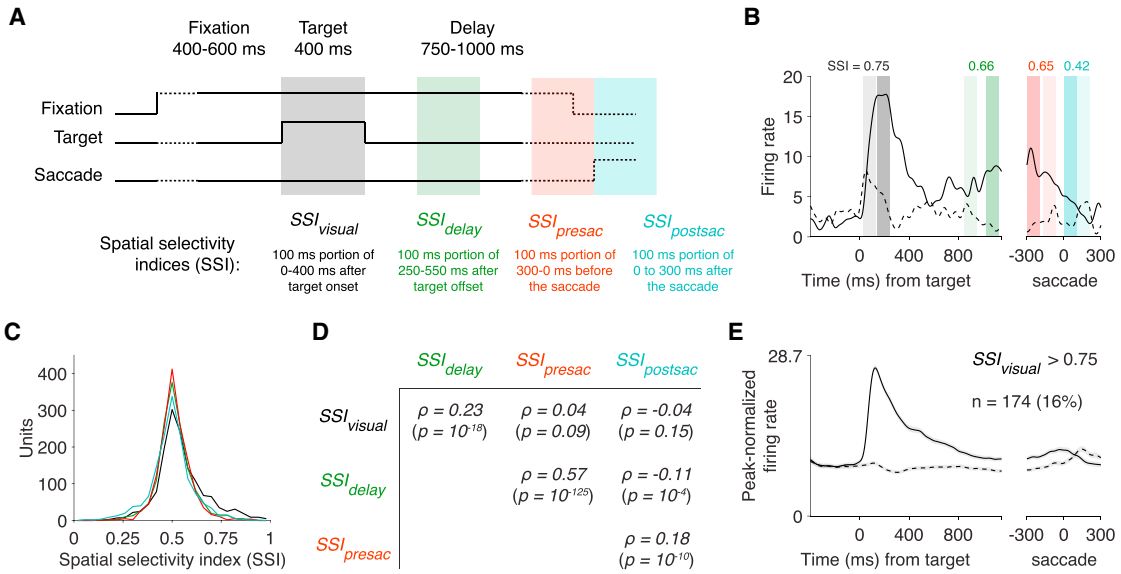


**Figure 3. A Memory-Guided Saccade Task Was Used to Index Spatially Selective**

(A) After acquiring fixation, a target (0.4° white square) appeared for 400 ms in either the left or the right hemifield. After a delay of 750–1,000 ms, the fixation point disappeared to cue animal to saccade to the remembered location of the target. A spatial selectivity index (SSI) was computed for each neuron during each of four task periods. The 100-ms time window with the maximal firing rate was selected separately for events in the contralateral and ipsilateral hemifield. Contralateral and ipsilateral firing rates were then compared using a receiver operating characteristic (ROC) analysis, and the resultant area under the ROC curve was the SSI. SSI > 0.5 indicates stronger activity at the contralateral location. Shaded regions indicate the range of time windows that can be selected for computing SSI.

(B) The peri-event time histogram (PETH) of an example unit was computed by convolving spikes with a Gaussian with $\sigma = 25$ ms. Solid lines correspond to trials with a contralateral target and saccade. Shaded bars represent the time windows used to compute the SSI during each period. Darker shading corresponds to contralateral events and lighter to ipsilateral events. The example was selected by picking the unit that had the lowest $SSI_{visual}$ among those with $SSI_{visual} > 0.75$.

(C) Distributions of SSI were centered close to 0.5, but $SSI_{visual}$ was skewed more toward 1 than other SSIs. N = 1,088.

(D) Partial correlations between each pair of SSI's. No reliable correlation was detected between $SSI_{visual}$ and either $SSI_{presac}$ or $SSI_{postsac}$. N = 1,088.

(E) Contra-selective visual units with $SSI_{visual} > 0.75$. To compute the population-averaged PETH, each unit was normalized by the maximal value in its PETH. The maximum of the y axis is 0.75 times the average peak firing rate across the population. Shading indicates mean ± SEM.

delay period ($SSI_{delay}$), before the saccade ($SSI_{presac}$), and after the saccade ($SSI_{postsac}$; Figures 3A and 3B). The SSI was equal to the area under the curve of a receiver operating characteristic analysis that classified the contralateral and ipsilateral responses during a given trial epoch. An SSI > 0.5 indicates stronger activity associated with the contralateral target or saccade than the corresponding ipsilateral event.

The distribution of $SSI_{visual}$ was modestly skewed toward 1 (median = 0.526; $p = 10^{-37}$; Figure 3C), indicating that more neurons in LPFC prefer contralateral stimuli than ipsilateral stimuli, as observed previously (Funahashi et al., 1990; Lennert and Martinez-Trujillo, 2011). The distributions were also greater than 0.5, though very modestly, for $SSI_{delay}$ (median = 0.505, $p = 10^{-5}$) and $SSI_{presac}$ (median = 0.504, $p = 10^{-4}$). $SSI_{visual}$ was larger than all other SSI (maximum $p = 10^{-10}$ across pairwise comparisons; firing rates were distribution-matched using bin sizes of 1 Hz). These results indicate that neurons in LPFC show stronger contralateral preference in their visual responses than in their delay-related or saccade-related responses.

No reliable partial correlation was observed between $SSI_{visual}$ and either $SSI_{presac}$ or $SSI_{postsac}$ (Figure 3D). A previous study reported overlap between visual receptive fields and presaccadic movement fields in LPFC (Bullock et al., 2017), which predicts a correlation between visual and presaccadic selectivity. Similarly, when we examined the correlation between $SSI_{visual}$ and $SSI_{presac}$ without controlling for selectivity in the delay period, we found a significant though modest correlation ($\rho = 0.22$, $p = 10^{-15}$). Therefore, in our dataset, the correlation between visual selectivity and presaccadic selectivity was largely due to correlations between visual and delay-related selectivity and between delay-related and presaccadic selectivity.

For subsequent analyses, we use as an example the subset of units with reliable contralateral preference in their visual responses, with $SSI_{visual} > 0.75$ (Figure 3E), but we also show that results generalize for a broad range of $SSI_{visual}$ thresholds for contralateral neurons. We focus on contralateral-preferring visual neurons because they outnumber ipsilateral-preferring cells and also because only 33/48 sessions had achieved behavioral isolation at ipsilateral location (Figure S1B). No difference in SSI was observed between neurons recorded during sessions isolating $\Delta c$ and sessions isolating $\Delta d'$ (Wilcoxon rank-sum test: maximum across different types of SSI was p = 0.39).

## Firing Rates Were Modulated When Subjects Changed either Their Criterion or Their Sensitivity

Figure 4A shows peri-event time histograms (PETH) of visual, contra-selective neurons ($SSI_{visual} > 0.75$) aligned to three trial events: sample onset, test onset, and saccade. These neurons robustly increased their responses to the sample stimulus when the animal either lowered its criterion or elevated its

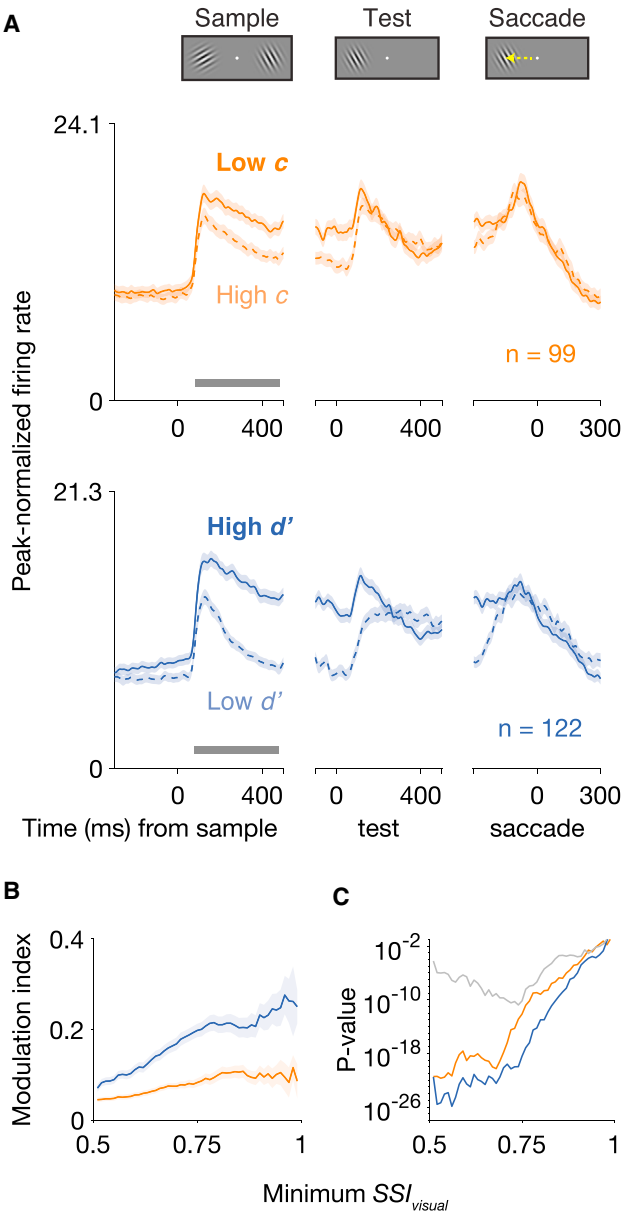


**Figure 4. Visual Responses Increased When Subjects either Lowered Their Criterion or Increased Their Sensitivity**

(A) PETH of units with $SSI_{visual}$ > 0.75 during the attention task. Orange lines represent units recorded during $\Delta c$ isolations, and blue lines represent units from $\Delta d'$ isolations. Spikes were convolved with a Gaussian with $\sigma$ = 8 ms. The maxima of the y axes are 0.75 times the average peak firing rate across the population. Shading indicates mean ± SEM. The bar near the x axis indicates the sample period (80–480 ms after sample onset). Either lowering *c* or elevating *d'* increased the firing rates of visual neurons throughout the sample period and into the test period.

(B) Modulation index (MI) of firing rates during the period 80–480 ms after sample onset as function of the minimum $SSI_{visual}$ of the units included for analysis. A $\Delta c$-related modulation index > 0 indicates stronger firing rates when the animal responded more frequently to the contralateral location. A $\Delta d'$-related MI > 0 indicates stronger firing rates when the animal was more accurate in discriminating between an orientation change and a match at the contralateral position. Shading indicates mean ± SEM across units. Modulation indices related to either $\Delta c$ or $\Delta d'$ are greater than 0 regardless of the $SSI_{visual}$ threshold. The sample sizes for minimum $SSI_{visual}$ = 0.51 were 492 units from $\Delta c$ isolations and 567 units from $\Delta d'$ isolations.

(C) p values were computed using the Wilcoxon signed-rank test for a zero median in the distribution of $\Delta c$-related modulation indices (orange) or $\Delta d'$-related indices (blue) and from the Wilcoxon rank-sum test of $\Delta c$- and $\Delta d'$-related indices having the same median (gray).

See also Figure S4.

sensitivity. To quantify the modulations in firing rates, we computed an index based on spike counts 80–480 ms after the onset of the sample (STAR Methods). The sample period was the most informative for assessing attention-related modulations because no behavioral decision or eye movement had yet occurred during this period. The modulation index associated with $\Delta c$, $MI_{\Delta c}$, was defined such that $MI_{\Delta c} > 0$ indicated a stronger response when the animal responded more frequently to the contralateral location. The modulation index associated with $\Delta d'$, $MI_{\Delta d'}$, was defined such that $MI_{\Delta d'} > 0$ indicated a stronger response when the animal was more accurate in discriminating between orientation changes and matches at the contralateral position.

We found the modulation indices related to either $\Delta c$ or $\Delta d'$ were substantially and reliably greater than zero (Figures 4B and 4C). The reliability of these effects held for a broad range of $SSI_{visual}$ thresholds. Modulations related to $\Delta d'$ were stronger than modulations related to $\Delta c$, but the two types of modulations were qualitatively similar, indicating that previously documented modulations in LPFC could reflect either behavioral change. These results indicate that visual neurons in LPFC encode robust signals related to behavioral changes in either criterion or sensitivity.

We next examined how attention-related modulations depend on other spatial selectivity indices. The trial-averaged firing rates of all units during the attention task were computed in nonoverlapping 200-ms bins. For each bin, we computed the partial correlation between the firing rate MI and each of $SSI_{visual}$, $SSI_{delay}$, $SSI_{presac}$, and $SSI_{postsac}$, while controlling for the other three SSIs (Figure 5). As expected from the previous analysis, there was reliable partial correlation between $SSI_{visual}$ and either $\Delta c$- or $\Delta d'$-related MI during the sample period. This correlation became negative during the test period and after the saccade, indicating that attention-related modulations can reverse sign after the animal completed a decision. Spatial selectivity during the delay ($SSI_{delay}$) correlated modestly with the attention-related modulations during the test period. Also, presaccadic spatial selectivity ($SSI_{presac}$) correlated with attention-related modulation in the sample period. These findings indicate that attention-related modulations depend on spatial selectivity measured in multiple different periods of the memory-guided saccade task, though the dependency appears to be strongest for the period after target presentation ($SSI_{visual}$).

Lastly, we measured the extent to which firing rate modulations during the sample period (80–480 ms after sample onset) can be explained using the SSIs. We fitted linear models to $\Delta c$-related and $\Delta d'$-related firing rates modulations during the sample period and selected the best model for each type of modulation using the Bayesian information criterion (Table S1). The proportion of variance explained by the optimal model

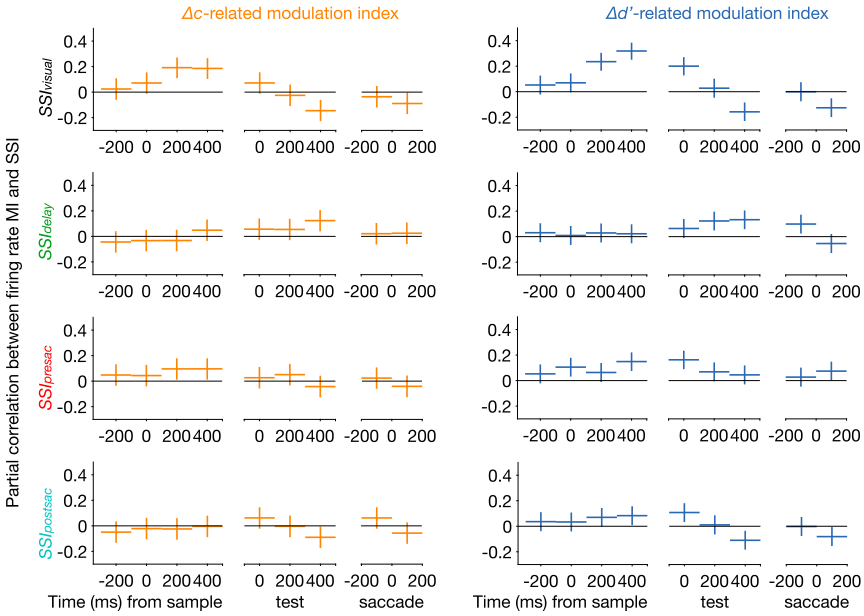


**Figure 5. Firing Rate Modulations Were Correlated with Spatial Selectivity Measured from Multiple Periods of the Memory-Guided Saccade Task**

Partial correlation between firing rate MI and each SSI (among $SSI_{visual}$, $SSI_{delay}$, $SSI_{presac}$, and $SSI_{postsac}$) while controlling for the other three SSI. The left column of panels (orange) corresponds to $\Delta c$-related modulations and the right column (blue) to $\Delta d'$-related modulations. Only the 33/48 sessions in which behavior were isolated at both the ipsilateral and contralateral locations were included for this analysis. n = 559 units from $\Delta c$ isolations and 667 units from $\Delta d'$ isolations. Vertical bars represent the 95% confidence interval.
See also Table S1.

was ~18% for $\Delta d'$-related modulations and ~13% for $\Delta c$-related modulations. The optimal models had a small intercept term ($MI_{\Delta c}$ = 0.14, p = $10^{-4}$, $MI_{\Delta d'}$ = 0.01, p = $10^{-21}$, t test), indicating that neurons in LPFC show lateralized $\Delta c$-related and $\Delta d'$-related modulation even if they display no spatial selectivity that could be detected in a memory-guided saccade task. We confirmed this to be the case by selecting neurons with $SSI_{visual}$, $SSI_{delay}$, $SSI_{presac}$, and $SSI_{postsac}$ all within the interval of [0.4, 0.6] and subsampling this population until the population average for each of all four SSIs was within [0.49, 0.51]. This nonselective population indeed showed small but significant modulation (mean $MI_{\Delta c}$ = 0.016, p = 0.05; mean $MI_{\Delta d'}$ = 0.019, p = 0.008; Wilcoxon signed-rank test).

## Pairwise Correlation and Fano Factor Were Also Modulated in Association with Either Behavioral Change

In addition to changes in firing rates, other attention-related neuronal modulations in visual cortex were found to be associated with behavioral changes in only sensitivity, but not criterion (Luo and Maunsell, 2015). These include reductions in spike count correlation between pairs of simultaneously recorded neurons (pairwise correlation) and across-trial variability in a neuron's firing rate (indexed as Fano factor). Pairwise correlation in prefrontal cortex has been used to decode the animal's trial-by-trial attentional state (Tremblay et al., 2015) and spatial working memory (Leavitt et al., 2017), but an attention-related reduction in correlation has not yet been reported in prefrontal cortex. Attention-related reductions in Fano factor were not observed in prefrontal cortex despite robust attention-related increases in firing rate, but these investigations focused on the frontal eye fields and not LPFC (Chang et al., 2012; Purcell et al., 2012).

Here, we found that either lowering criterion or elevating sensitivity was associated with a reduction in both pairwise correlation and Fano factor among contra-selective visual neurons during the sample period (Figure 6). The information in the neural population for discriminating between the two orientations, measured as linear Fisher information (Kanitscheider et al., 2015), increased modestly in association with either a decrease in criterion or an increase in sensitivity, though these changes in population coding did not depend on changes in pairwise correlation (Figure S5). The changes in pairwise correlation and Fano factor further support the idea that robust neuronal signals associated with either component of visuospatial attention are present in LPFC.

## Reliable $\Delta c$-Related Modulations Were Observed in the Population Activity in LPFC, but Not in Area V4

In a previous experiment (Luo and Maunsell, 2015), neuronal responses were recorded from area V4 in visual cortex while behavioral changes in either criterion or sensitivity were isolated using the same task as used here. We therefore directly compared the modulations in the activity of visual neurons between the two areas (Figure 7A). Both areas showed robust $\Delta d'$-related modulations in firing rates, pairwise correlation, and Fano factor. However, whereas LPFC had reliable $\Delta c$-related modulations, the $\Delta c$-related modulations in V4 in firing rate and Fano factor were weak. Moreover, the $\Delta c$-related modulations in pairwise correlation in V4 were opposite in sign to usual attention-related changes in correlation (Cohen and Maunsell, 2009; Mitchell et al., 2009).

It appears that neuronal responses in V4 are associated with only one component of visuospatial attention, while LPFC is associated with both. But it might instead be that $\Delta c$-related modulations were detectable in LPFC only because attentional modulations are generally stronger there, continuing the increase of attentional modulation with the level of visual cortical hierarchy (Maunsell and Cook, 2002). To evaluate these two possibilities, we fitted two models to the $\Delta c$- and $\Delta d'$-related modulations in the two brain areas (STAR Methods). In model 1, the mean of $\Delta c$-related changes across neurons in V4 was set to zero, while the population-means of $\Delta d'$-related modulations in V4 and of $\Delta c$- and $\Delta d'$-related modulations in LPFC were free to vary. In model 2, two parameters were assigned to the

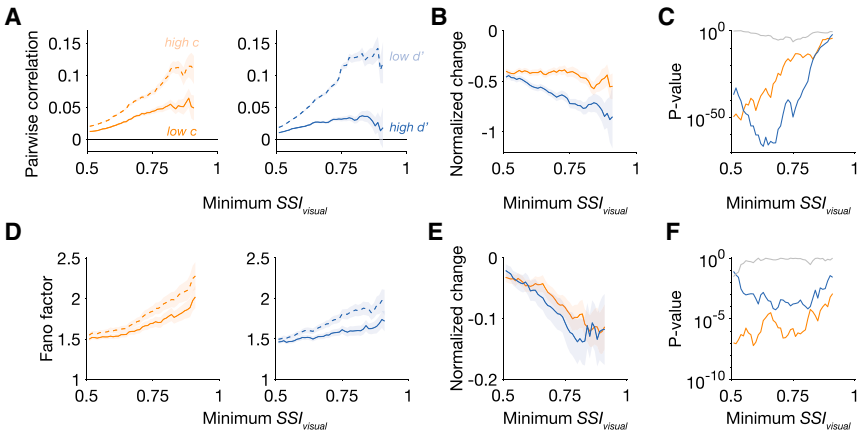


**Figure 6. Pairwise Correlation and Fano Factor Decreased When the Animal either Lowered Its Criterion or Elevated Its Sensitivity**

(A) Pairwise correlations between contra-selective visual units decreased when the animal performed with a lower criterion or a higher sensitivity at the contralateral hemifield. Correlations were computed using firing rates during the sample period (80–480 ms after sample onset). The sample sizes for minimum $SSI_{visual}$ = 0.51 and 0.75 were 25,273 and 988 pairs from $\Delta c$ isolations and 26,635 and 974 pairs from $\Delta d'$ isolations.

(B) Changes in pairwise correlation were normalized by the population-averaged correlation in either *high c contra* condition (for $\Delta c$ isolations) or the *low d' contra* condition (for $\Delta d'$ isolations).

(C) p values were computed using the Wilcoxon signed-rank test for a zero median in the distribution of normalized changes related to either $\Delta c$ (orange) or $\Delta d'$ (blue) and using the Wilcoxon rank-sum test of $\Delta c$- and $\Delta d'$-related correlation changes having the same median (gray).

(D) Fano factor of contra-selective visual neurons also decreased when the animal performed with either a lower criterion or higher sensitivity. Fano factor was computed using spike counts during the sample period. The sample sizes for $SSI_{visual}$ = 0.51 and 0.75 were 492 and 99 units from $\Delta c$ isolations and 567 and 122 units from $\Delta d'$ isolations.

(E) Normalized change in Fano factor.

(F) p values for normalized changes in Fano factor.

See also Figure S5.

population means of $\Delta c$- and $\Delta d'$-related modulations, and a third parameter determined the multiplicative relationship between the population-averaged modulations in V4 and LPFC. The two models were fitted using maximum likelihood estimation. Because the two models had the same number of parameters, they were compared using their maximized likelihoods (Figures 7B and 7C). Across a broad range of $SSI_{visual}$ thresholds, the likelihood of model 1 was significantly higher than that of model 2 for firing rate and pairwise correlation. For Fano factor, the likelihood of model 1 was also larger, but it was significantly larger within only a limited range of $SSI_{visual}$ thresholds, perhaps due to the typically smaller effect sizes of Fano factor modulations. This analysis indicates that the difference in attention-related changes between V4 and LPFC is not merely a stronger modulation in the latter, but that LPFC provides signals associated with a component of attention that were absent in population-averaged activity in V4.

However, individual neurons in V4 were modulated in association with $\Delta c$. A greater than expected fraction of V4 neurons significantly changed their firing rates in association with changes in behavioral criterion (6.5% of units at the p = 0.01 threshold), even though firing rates across these neurons averaged to near zero (Figures 7D and 7E). Therefore, when a subject changes its criterion, while firing rate changes are not detected in the population-averaged activity in V4, a small fraction of individual neurons significantly changed their firing rates.

## Modulations Related to Criterion and Sensitivity Changes Were Correlated across LPFC Neurons

Firing rate modulations related to either $\Delta c$ or $\Delta d'$ were often observed for units recorded from the same electrodes (Figure S6), raising the question of the extent to which $\Delta c$- and $\Delta d'$-related modulations are correlated across LPFC neurons. A correlation would indicate that $\Delta c$- and $\Delta d'$-related modulations are combined within LPFC or in a common input, while a lack of correlation indicates that LPFC neurons receive independent $\Delta c$- and $\Delta d'$-related inputs. While we did not isolate $\Delta c$ and $\Delta d'$ on the same day, we can measure the correlation across units that were recorded in multiple sessions and identified by the tracking algorithm to be putatively the same neuron.

Figure 8A plots the MIs of each putative unit with a $SSI_{visual}$ > 0.75 from each pair of $\Delta c$-isolation and $\Delta d'$-isolation sessions in which that unit was recorded. Putative units showed a strong correlation between the two modulations, with the $\Delta d'$-related MI averaging about 1.7 times the $\Delta c$-related MI. The correlation suggests that $\Delta c$- and $\Delta d'$-related signals are at least partially combined within LPFC or have partially overlapping inputs.

The measured correlation between $\Delta c$- and $\Delta d'$- related MIs should be considered a lower-bound, owing to noise arising from finite trial counts, error in tracking units across days, and uncontrolled behavioral difference between sessions. To assess how much these factors reduce the measured correlation, we sought to normalize the $\Delta c/\Delta d'$ correlation by the same-isolation correlation, i.e., the correlation between MIs related to the same behavioral change, which we assume to approximate the maximum correlation that can be extracted given the various measurement errors (STAR Methods).

Figure 8B plots the MIs from each pair of sessions with the same behavioral isolation for each putative unit with a $SSI_{visual}$ > 0.75 that was recorded during both of those sessions. The correlation between same-isolation pairs was similar to that measured for $\Delta c/\Delta d'$ pairs, suggesting that the correlation in Figure 8A is about as strong as can be expected given the various sources of measurement error. Across a broad range

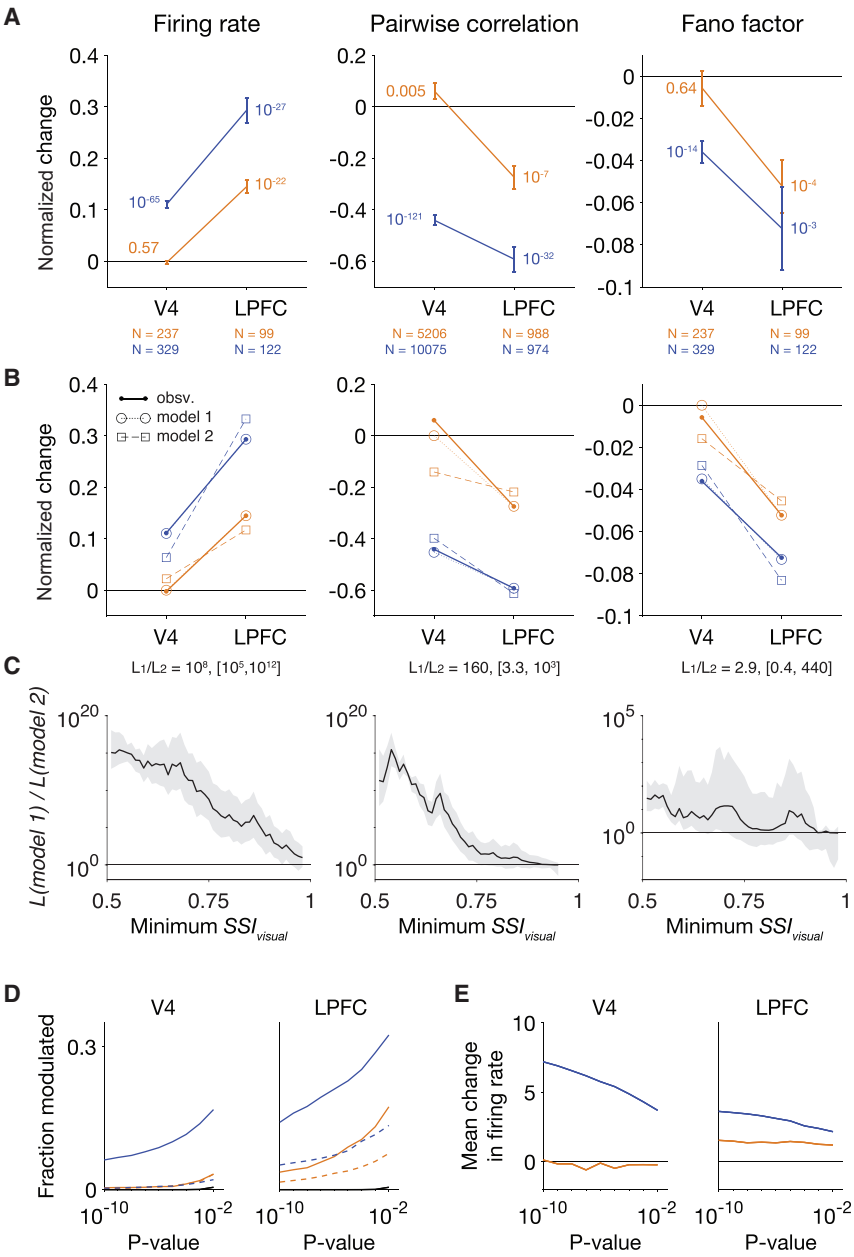


**Figure 7. Neuronal Modulations Related to $\Delta c$ Were Reliably Observed in the Population Activity of LPFC, but Not of Area V4**

(A) Normalized change in firing rate, pairwise correlation, and Fano factor of neurons in LPFC and V4 with $SSI_{visual}$ > 0.75. Error bars represent mean ± SEM. p values were computed using a Wilcoxon signed-rank test.

(B) Maximum likelihood fits from model 1 (circles and dotted line) and from model 2 (squares and dashed lines). Model 1 assumes zero $\Delta c$-related modulations in V4, and model 2 assumes a multiplicative relationship between the modulations in V4 and LPFC. Each model has three parameters. The likelihood ratio and its bootstrapped 95% confidence intervals are indicated at the bottom of the panel.

(C) The likelihood ratio of model 1 to model 2 across thresholds of $SSI_{visual}$. The likelihood ratio was greater than one (favoring model 1) for all thresholds of $SSI_{visual}$ for all three measures. For firing rate and pairwise correlation, the bootstrapped 95% confidence interval (shading) of the likelihood ratio was greater than one for a broad range of $SSI_{visual}$.

(D) The fraction of neurons whose firing rates were significantly modulated in association with either $\Delta c$ (orange lines) or $\Delta d'$ (blue lines) plotted against the threshold for statistical significance. Solid lines indicate the fraction with significantly increased firing rates, and dashed lines indicate the fraction with a significant decrease. Black lines indicate the discovery rate at chance (p value divided by two).

(E) The mean firing rate change across all neurons with an individually significant increase or decrease in their firing rates.

of minimum $SSI_{visual}$, the $\Delta c/\Delta d'$ correlation was statistically indistinguishable from the same-isolation correlation (Figure 8C). Normalized by the same-isolation correlation, the $\Delta c/\Delta d'$ correlation was indistinguishable from one (Figure 8D). This result is consistent with LPFC neurons receiving a common input that combines $\Delta c$- and $\Delta d'$-related modulations or that separate $\Delta c$- and $\Delta d'$-related inputs are combined within LPFC.

In contrast to LPFC, the normalized $\Delta c/\Delta d'$ correlation within V4 pairs was indistinguishable from zero (slightly negative if a more liberal statistical criterion than p = 0.05 were used; Figure 8D). The normalized $\Delta c/\Delta d'$ correlations were reliably different between LPFC and V4 across thresholds of minimum $SSI_{visual}$. These findings support the idea that the attention-related modulations in LPFC and V4 are part of distinct mechanisms.

These results were verified by two additional analyses. First, because MI pairs sampled from different isolations were on average more separated in time than MI pairs sampled from the same isolation (Figure S7A), the correlations shown in Figures 8C and 8D were computed after matching the distributions of the number of intervening sessions between $\Delta c/\Delta d'$ pairs and same-isolation pairs (Churchland et al., 2010) (Figure S7B). The same analyses performed without distribution-matching showed qualitatively similar results (Figures S7C and S7D). A second analysis in which we computed correlations across units rather than pairs of MIs also showed a reliable difference in $\Delta c/\Delta d'$ correlations between LPFC and V4 (Figure S7E).

## DISCUSSION

A challenge in understanding visuospatial attention involves partitioning and organizing its related neuronal signals across the brain. Classical signal detection theory is useful for this purpose

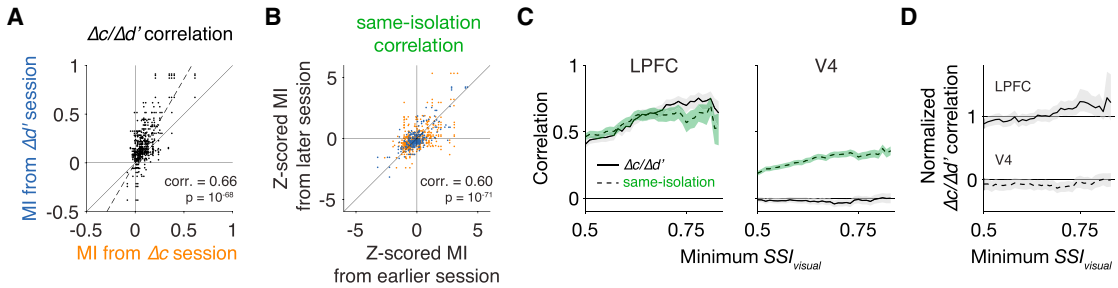


**Figure 8. Firing Rate Modulations Related to *Δc* and *Δd'* Were Correlated across Neurons in LPFC**

(A) Firing rate modulation indices (MI) from each pair of *Δc*-isolation session and *Δd'*-isolation session from each unit that was recorded in both sessions. Each marker indicates one pair of MI (576 pairs from 47 units that had an $SSI_{visual} > 0.75$). Both MIs of each pair belonged to the same unit. The slope of the regressed line (dashed, no intercept) was 1.72.

(B) MIs from each pair of sessions with the same isolation from each unit that was recorded in both sessions. To combine MI pairs from *Δc* sessions (orange) and from *Δd'* sessions (blue), MIs of each isolation type were *Z* scored (710 pairs from 79 units that had an $SSI_{visual} > 0.75$).

(C) Correlations between *Δc*/*Δd'* MI pairs and between same-isolation MI pairs. Correlations were calculated after distribution-matching the number of intervening sessions between each pair (Figures S7A and S7B). Shading indicates 95% confidence interval.

(D) Normalized *Δc*/*Δd'* correlations were computed by dividing *Δc*/*Δd'* correlations by same-isolation correlations, and 95% confidence intervals were computed using a bootstrap procedure. Normalized correlations in LPFC were statistically indistinguishable from 1 and reliably different from normalized correlations in V4, which were indistinguishable from 0.

See also Figure S7.

because it accurately describes performance in simple perceptual tasks using two orthogonal parameters (Green and Swets, 1966; Macmillan and Creelman, 2004). We hypothesized that the brain areas related with visuospatial attention can be distinguished according to whether the neuronal modulations therein correspond to behavioral changes in criterion, sensitivity, or both. This hypothesis was suggested by our previous result from visual cortex showing that attention-related signals in area V4 were associated with only changes in the subject's sensitivity but not its criterion (Luo and Maunsell, 2015). This previous result suggests that other brain areas besides V4 underlie changes in the subject's criterion. However, it was unclear what kind of neuronal signals are related to attention-related criterion changes.

We found that the responses of visual neurons in LPFC were robustly modulated in association with changes in the subject's criterion. Criterion-related modulations were qualitatively similar to sensitivity-related modulations, and the two modulations were correlated on a cell-by-cell basis. These results are qualitatively different from previous results in area V4, therefore indicating that attention-related modulations in different brain areas can be meaningfully distinguished according to whether they correspond to behavioral changes in the subject's sensitivity, criterion, or both. This scheme provides a step forward from treating neuronal changes related to attention as a monolithic signal, as has been often done in neurophysiological studies of visuospatial attention.

## Is Criterion Change a Component of Visuospatial Attention?

Changes in an observer's criterion are consistent with an aspect shared by many definitions of attention: the "selection" of a stimulus to be perceived in greater detail, held in working memory, or acted on (Carrasco, 2011; Desimone and Duncan, 1995; Knudsen, 2007). This aspect of attention is consistent with the way criterion changes are operationalized here: when a subject lowers its criterion at a location, the subject "selects" the stimulus to be a likely target for an oculomotor response.

Considering criterion changes to be a component of attention provides an approach for partitioning the neuronal signals across different areas. The distinction between bottom-up and top-down attention (Buschman and Miller, 2007) and between spatial and feature attention (Bichot et al., 2015) have been useful in differentiating the contributions of different brain areas. However, even within the domain of top-down spatial attention, a large number of brain areas have been implicated (Desimone and Duncan, 1995). The distinction between criterion and sensitivity allows these areas to be functionally differentiated rather than viewed as homogeneous.

There is also a practical advantage in treating criterion change as a component of attention: many of the tasks that documented the neuronal signals related to visuospatial attention were designed such that the subject could have changed its criterion to improve its performance (Luo and Maunsell, 2015; Sridharan et al., 2017). To exclude behavioral criterion as a component of attention implies excluding or raising serious doubts regarding many studies included in the neurophysiological literature of attention. The potential contribution of changes in criterion or other decisional processes to the effects of attention is pervasive. The behavioral improvements in human observers performing the Posner cuing paradigm could depend entirely on decisional changes (Eckstein et al., 2002). The attention-related effects in not only detection tasks but also discrimination tasks could depend entirely on changes in criterion (Sridharan et al., 2017). Many current definitions of attention in the neurophysiological literature include (but do not necessarily emphasize) the decisional and oculomotor processes on which behavioral criterion likely depend (Desimone and Duncan, 1995; Knudsen, 2007; Krauzlis et al., 2013; Moore and Zirnsak, 2017).

## Attention-Related Neuronal Changes in Prefrontal Cortex

We found that firing rate modulations related to the two behavioral changes were highly correlated on a cell-by-cell basis. This finding suggests that $\Delta c$- and $\Delta d'$-related modulations arrive in a common signal in LPFC. However, it is also possible that $\Delta c$- and $\Delta d'$-related signals have separate origins, and that a mechanism within LPFC restricts the extent to which individual cells are modulated by either signal. Specifically, the variability in the size of attention-related modulations of firing rates across neurons in visual cortex are well explained by the strength of response normalization (Boynton, 2009; Lee and Maunsell, 2009; Ni et al., 2012; Reynolds and Heeger, 2009). The correlation in $\Delta c$- and $\Delta d'$-related modulations across neurons in LPFC might arise in part because both modulations depend on normalization mechanisms within LPFC.

This finding further raises the question of the extent to which the modulations in LPFC make up a generic attentional signal that does not distinguish between whether the animal has lowered its criterion, increased its sensitivity, or done both. A prominent view is that the firing rates of visual neurons in sensorimotor areas of the brain, such as the LPFC, frontal eye fields, and lateral intraparietal cortex, represents a generic "priority" or "salience" signal that can be used either guide a saccade or to allocate attention (Thompson and Bichot, 2005; Bisley and Goldberg, 2010). Our results indicate that the modulations in LPFC can be interpreted as a generic signal of this sort, and that sensitivity changes might be more closely coupled with such a signal than are criterion changes and are therefore associated with stronger modulations. However, further studies could reveal distinctions that make the notion of any generic signal untenable.

We observed reductions in the Fano factor and pairwise correlation of visual neurons when the animal changed either its criterion or its sensitivity. These two changes in spiking variability have been consistently observed in visual cortex when attention was manipulated (Cohen and Maunsell, 2009; Mitchell et al., 2007, 2009; Luo and Maunsell, 2015). In the FEF, attention-related Fano factor changes were not observed despite robust increases in firing rates (Chang et al., 2012; Purcell et al., 2012). While it is conceivable that the absence of Fano factor changes indicates different mechanisms underlying attentional changes in FEF and LPFC, reductions in Fano factor have small effect sizes. Even in visual cortex, changes in Fano factor may go undetected when increases in firing rate are readily apparent (McAdams and Maunsell, 1999). Attention-related decreases in pairwise correlation have not been reported in previous attention studies that performed multi-electrode recording in LPFC (e.g., Kadohisa et al., 2013; Tremblay et al., 2015). The association between criterion changes with changes in pairwise correlation is interesting because changes in pairwise correlation have been postulated as a mechanism for enhancing the animal's perceptual sensitivity (Cohen and Maunsell, 2009; Mitchell et al., 2009). Our finding indicates that changes in pairwise correlation can occur even when there is no change in the animal's perceptual sensitivity. A recent study indicates that attention-related changes in pairwise correlation in visual cortex can arise from response normalization mechanisms (Verhoef and Maunsell, 2017). Both changes in firing rate and in pairwise correlation related to behavioral changes in both criterion and sensitivity could depend to some extent on normalization mechanisms.

## Brain Structures Associated with Attentional Changes in Criterion and Sensitivity

We found that the attention-related modulations in LPFC are associated with changes in both the subject's criterion and its sensitivity, whereas such modulations in area V4 of visual cortex are associated only with sensitivity changes. Two recent studies examined whether the superior colliculus (SC) contributes to attention through changes in criterion or sensitivity (Sridharan et al., 2017; Lovejoy and Krauzlis, 2017). Both studies applied a multidimensional extension to classical signal detection theory. While the definitions of criterion are not identical across these studies (and our study), they represent likely overlapping neural processes. Together these two studies indicate that SC contributes to attention through changes in both criterion and sensitivity, but with criterion being the dominant contribution.

Sridharan et al. (2017) analyzed four previous studies in which the SC was perturbed, through either microstimulation or pharmacological inactivation, while subjects performed tasks that required spatial attention (Cavanaugh and Wurtz, 2004; Lovejoy and Krauzlis, 2010; Müller et al., 2005; Zénon and Krauzlis, 2012). For all the studies, the changes in behavioral performance resulting from collicular perturbation could be explained by a change in criterion without a change in sensitivity. Furthermore, for two studies in which models were directly fitted to the data, criterion changes alone (Zénon and Krauzlis, 2012) and a combination of changes in criterion and sensitivity (Lovejoy and Krauzlis, 2010) best accounted for the data. The analyses in Sridharan et al., 2017 therefore suggest that the predominant contribution of SC to spatial attention involves changing the subject's criterion across different visual locations.

Lovejoy and Krauzlis (2017) trained monkeys to discriminate the direction of a moving stimulus that could appear at a random one of four locations (uncued condition) or appear at a cued location (cued condition). The effect of attention on performance was defined as the increase in $d'$ between the uncued and the cued conditions. When SC was pharmacologically inactivated, the cue-induced increase in $d'$ could not be detected in the affected region of visual space, indicating that SC is necessary for attention-related enhancement in behavioral sensitivity. During inactivation, the subject's criterion was also affected by the inactivation such that subjects were less inclined to saccade to the affected region.

Given that both LPFC and SC are associated with both criterion changes and sensitivity changes, it is possible that no brain structure is entirely associated with only criterion changes. Some structures, such as V4 and LPFC, may be more strongly modulated in association with sensitivity changes than criterion changes, while others, possibly SC, are more closely linked to criterion changes. If this were the case, when animals are trained to change either only their criterion or only their sensitivity, overlapping groups of brain structures should be modulated, but to different extents.

## Conclusion

Distinguishing and organizing the widespread attention-related signals in the brain is necessary for a deeper understanding of

visuospatial attention. Our results indicate that neuronal changes in different brain areas related to attention can be partitioned using the indices of criterion and sensitivity from signal detection theory. Future work capitalizing on the distinction between criterion and sensitivity will likely provide insights into the mechanisms of visuospatial attention.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Criterion (c) and sensitivity (d')
  - Attention task
  - Memory-guided saccade task
  - Tracking neurons across days
  - Spatial selectivity index (SSI)
  - Modulation index (MI) and peri-event time histogram (PETH)
  - Partial correlation between MI and SSI
  - Pairwise correlation and Fano factor
  - Comparison between LPFC and V4
  - Correlation between Δc- and Δd'-related firing rate MI
  - Linear Fisher information
- QUANTIFICATION AND STATISTICAL ANALYSES
- DATA AND SOFTWARE AVAILABILITY

### SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and one table and can be found with this article online at https://doi.org/10.1016/j.neuron.2018.02.007.

### ACKNOWLEDGMENTS

We thank Richard T. Born, Jackson J. Cone, Supriya Ghosh, and Bram-Ernst Verhoef for comments and discussion; Peter J. Wiese for technical assistance; Michael Petrides for advice on the anatomical locations of implanted arrays; and Ruben Coen-Cagli on advice on computing Fisher information. This work was supported by grants from the NIH (R01EY005911 and F31MH103895).

### AUTHOR CONTRIBUTIONS

T.Z.L. and J.H.R.M. designed the experiments, performed the surgeries, and wrote the paper. T.Z.L. performed the experiments and analyzed the data.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 28, 2017
Revised: January 17, 2018
Accepted: February 4, 2018
Published: March 1, 2018

### SUPPORTING CITATIONS

The following references appear in the Supplemental Information: Petrides and Pandya (2002).

### REFERENCES

Arcizet, F., Mirpour, K., Foster, D.J., and Bisley, J.W. (2017). Activity in LIP, but not V4, matches performance when attention is spread. Cereb. Cortex. Published online October 23, 2017. https://doi.org/10.1093/cercor/bhx274.

Bashinski, H.S., and Bacharach, V.R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. Percept. Psychophys. *28*, 241–248.

Bichot, N.P., Heard, M.T., DeGennaro, E.M., and Desimone, R. (2015). A source for feature-based attention in the prefrontal cortex. Neuron *88*, 832–844.

Bisley, J.W., and Goldberg, M.E. (2010). Attention, intention, and priority in the parietal lobe. Annu. Rev. Neurosci. *33*, 1–21.

Boch, R.A., and Goldberg, M.E. (1989). Participation of prefrontal neurons in the preparation of visually guided eye movements in the rhesus monkey. J. Neurophysiol. *61*, 1064–1084.

Boussaoud, D., and Wise, S.P. (1993). Primate frontal cortex: neuronal activity following attentional versus intentional cues. Exp. Brain. Res. *95*, 15–27.

Boynton, G.M. (2009). A framework for describing the effects of attention on visual responses. Vision Res. *49*, 1129–1143.

Briggs, F., Mangun, G.R., and Usrey, W.M. (2013). Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits. Nature *499*, 476–480.

Bruce, C.J., and Goldberg, M.E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. J. Neurophysiol. *53*, 603–635.

Bullock, K.R., Pieper, F., Sachs, A.J., and Martinez-Trujillo, J.C. (2017). Visual and presaccadic activity in area 8Ar of the macaque monkey lateral prefrontal cortex. J Neurophysiol *118*, 15–28.

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.

Carrasco, M. (2011). Visual attention: the past 25 years. Vision Res. *51*, 1484–1525.

Cavanaugh, J., and Wurtz, R.H. (2004). Subcortical modulation of attention counters change blindness. J. Neurosci. *24*, 11236–11243.

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat. Neurosci. *13*, 369–378.

Chang, M.H., Armstrong, K.M., and Moore, T. (2012). Dissociation of response variability from firing rate effects in frontal eye field neurons during visual stimulation, working memory, and attention. J. Neurosci. *32*, 2204–2216.

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

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. *18*, 193–222.

Dickey, A.S., Suminski, A., Amit, Y., and Hatsopoulos, N.G. (2009). Single-unit stability using chronically implanted multielectrode arrays. J. Neurophysiol. *102*, 1331–1339.

Downing, C.J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. J. Exp. Psychol. Hum. Percept. Perform. *14*, 188–202.

Eckstein, M.P., Shimozaki, S.S., and Abbey, C.K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. J. Vis. *2*, 25–45.

Fraser, G.W., and Schwartz, A.B. (2012). Recording from the same neurons chronically in motor cortex. J. Neurophysiol. *107*, 1970–1978.

Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1990). Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. J. Neurophysiol. *63*, 814–831.

Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1991). Neuronal activity related to saccadic eye movements in the monkey's dorsolateral prefrontal cortex. J. Neurophysiol. *65*, 1464–1483.

Goldberg, M.E., and Wurtz, R.H. (1972). Activity of superior colliculus in behaving monkey. II. Effect of attention on neuronal responses. J. Neurophysiol. *35*, 560–574.

Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics (Peninsula Publishing).

Gregoriou, G.G., Gotts, S.J., and Desimone, R. (2012). Cell-type-specific synchronization of neural activity in FEF with V4 during attention. Neuron *73*, 581–594.

Hawkins, H.L., Hillyard, S.A., Luck, S.J., Mouloua, M., Downing, C.J., and Woodward, D.P. (1990). Visual attention modulates signal detectability. J. Exp. Psychol. Hum. Percept. Perform. *16*, 802–811.

Ignashchenkova, A., Dicke, P.W., Haarmeier, T., and Thier, P. (2004). Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. Nat. Neurosci. *7*, 56–64.

Kadohisa, M., Petrov, P., Stokes, M., Sigala, N., Buckley, M., Gaffan, D., Kusunoki, M., and Duncan, J. (2013). Dynamic construction of a coherent attentional state in a prefrontal cell population. Neuron *80*, 235–246.

Kanitscheider, I., Coen-Cagli, R., Kohn, A., and Pouget, A. (2015). Measuring Fisher information accurately in correlated neural populations. PLoS Comput. Biol. *11*, e1004218.

Knudsen, E.I. (2007). Fundamental components of attention. Annu. Rev. Neurosci. *30*, 57–78.

Krauzlis, R.J., Lovejoy, L.P., and Zénon, A. (2013). Superior colliculus and visual spatial attention. Annu. Rev. Neurosci. *36*, 165–182.

Leavitt, M.L., Pieper, F., Sachs, A.J., and Martinez-Trujillo, J.C. (2017). Correlated variability modifies working memory fidelity in primate prefrontal neuronal ensembles. Proc. Natl. Acad. Sci. USA *114*, E2494–E2503.

Lee, J., and Maunsell, J.H. (2009). A normalization model of attentional modulation of single unit responses. PLoS One *4*, e4651.

Lennert, T., and Martinez-Trujillo, J. (2011). Strength of response suppression to distracter stimuli determines attentional-filtering performance in primate prefrontal neurons. Neuron *70*, 141–152.

Lovejoy, L.P., and Krauzlis, R.J. (2010). Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments. Nat. Neurosci. *13*, 261–266.

Lovejoy, L.P., and Krauzlis, R.J. (2017). Changes in perceptual sensitivity related to spatial cues depends on subcortical activity. Proc. Natl. Acad. Sci. USA *114*, 6122–6126.

Luo, T.Z., and Maunsell, J.H.R. (2015). Neuronal modulations in visual cortex are associated with only one of multiple components of attention. Neuron *86*, 1182–1188.

Lynch, J.C., Mountcastle, V.B., Talbot, W.H., and Yin, T.C. (1977). Parietal lobe mechanisms for directed visual attention. J. Neurophysiol. *40*, 362–389.

Macmillan, N.A., and Creelman, D.C. (1990). Response bias: characteristics of detection theory, threshold theory, and "nonparametric" indexes. Psychol. Bull. *107*, 401–413.

Macmillan, N.A., and Creelman, D.C. (2004). Detection Theory: A User's Guide (Lawrence Erlbaum Associates).

Maunsell, J.H.R. (2015). Neuronal mechanisms of visual attention. Annu. Rev. Vis. Sci. *1*, 373–391.

Maunsell, J.H.R., and Cook, E.P. (2002). The role of attention in visual processing. Philos. Trans. R. Soc. Lond. B Biol. Sci. *357*, 1063–1072.

McAdams, C.J., and Maunsell, J.H. (1999). Effects of attention on the reliability of individual neurons in monkey visual cortex. Neuron *23*, 765–773.

McAlonan, K., Cavanaugh, J., and Wurtz, R.H. (2008). Guarding the gateway to cortex with attention in visual thalamus. Nature *456*, 391–394.

Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. Neuron *55*, 131–141.

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.

Moore, T., and Zirnsak, M. (2017). Neural mechanisms of selective visual attention. Annu. Rev. Psychol. *68*, 47–72.

Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. Science *229*, 782–784.

Müller, H.J., and Findlay, J.M. (1987). Sensitivity and criterion effects in the spatial cuing of visual attention. Percept. Psychophys. *42*, 383–399.

Müller, H.J., and Humphreys, G.W. (1991). Luminance-increment detection: capacity-limited or not? J. Exp. Psychol. Hum. Percept. Perform. *17*, 107–124.

Müller, J.R., Philiastides, M.G., and Newsome, W.T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. Proc. Natl. Acad. Sci. USA *102*, 524–529.

Ni, A.M., Ray, S., and Maunsell, J.H. (2012). Tuned normalization explains the size of attention modulations. Neuron *73*, 803–813.

Petrides, M., and Pandya, D.N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur. J. Neurosci. *11*, 1011–1036.

Petrides, M., and Pandya, D.N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. Eur. J. Neurosci *16*, 291–310.

Posner, M.I., Snyder, C.R., and Davidson, B.J. (1980). Attention and the detection of signals. J. Exp. Psychol. *109*, 160–174.

Purcell, B.A., Heitz, R.P., Cohen, J.Y., and Schall, J.D. (2012). Response variability of frontal eye field neurons modulates with sensory input and saccade preparation but not visual search salience. J. Neurophysiol. *108*, 2737–2750.

Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. Neuron *61*, 168–185.

Sridharan, D., Steinmetz, N.A., Moore, T., and Knudsen, E.I. (2017). Does the superior colliculus control perceptual sensitivity or choice bias during attention? Evidence from a multialternative decision framework. J. Neurosci. *37*, 480–511.

Thompson, K.G., and Bichot, N.P. (2005). A visual salience map in the primate frontal eye field. Prog. Brain Res. *147*, 251–262.

Thompson, K.G., Biscoe, K.L., and Sato, T.R. (2005). Neuronal basis of covert spatial attention in the frontal eye field. J. Neurosci. *25*, 9479–9487.

Tremblay, S., Pieper, F., Sachs, A., and Martinez-Trujillo, J. (2015). Attentional filtering of visual information by neuronal ensembles in the primate lateral prefrontal cortex. Neuron *85*, 202–215.

Verhoef, B.-E., and Maunsell, J.H.R. (2017). Attention-related changes in correlated neuronal activity arise from normalization mechanisms. Nat. Neurosci. *20*, 969–977.

Walker, A.E. (1940). A cytoarchitectural study of the prefrontal area of the macaque monkey. J. Comp. Neurol. *73*, 59–86.

Wurtz, R.H., and Mohler, C.W. (1976). Enhancement of visual responses in monkey striate cortex and frontal eye fields. J. Neurophysiol. *39*, 766–772.

Wyart, V., Nobre, A.C., and Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. Proc. Natl. Acad. Sci. USA *109*, 3593–3598.

Zénon, A., and Krauzlis, R.J. (2012). Attention deficits without cortical neuronal deficits. Nature *489*, 434–437.

# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Deposited Data | | |
| Raw data | This paper | Mendeley Data: https://doi.org/10.17632/7ns5zz87bp.1 |
| Software and Algorithms | | |
| MATLAB | MathWorks | RRID:SCR_001622 |
| Lablib | Maunsell Lab | https://github.com/MaunsellLab/Lablib-Public-05-July-2016 |
| Offline Sorter | Plexon | Offline Sorter v3.3.5 – Windows 7 (64 bit) |
| "Tracking neurons over multiple days" | MATLAB Central | 30113 |
| Other | | |
| Utah Array | Blackrock Microsystems | 6x8 CerePort array assembly |

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Thomas Zhihao Luo (thomas.zhihao.luo@gmail.com).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal use procedures were approved by The University of Chicago Institutional Care and Use Committee (IACUC) and carried out in accordance with National Institute of Health standards. Two male, eight-year-old, 9-11 kg rhesus macaques (*Macaca mulatta*) were used in the study. Access to water was scheduled to training or recording sessions that occurred 1-3 hours per day. Eye movements were tracked using a video system (EyeLink 1000, 500 Hz). Before training, each animal was implanted with a head post, and after training was completed, each animal was implanted with a pair of 6x8 Utah microelectrode arrays (Blackrock Microsystems) in lateral prefrontal cortex.

## METHOD DETAILS

### Criterion (c) and sensitivity (d')

Criterion was indexed as *criterion location* (*c*) as described in Macmillan and Creelman (2004):

$$c = \frac{1}{2}\left[\Phi^{-1}(\textit{hit rate}) + \Phi^{-1}(\textit{false alarm rate})\right].$$

$\Phi^{-1}$ is the inverse normal cumulative distribution function. When $c = 0$, the subject shows no bias toward reporting either targets or nontargets. In the signal detection model (Figure 1A), this is the value on the x axis where the two normal distributions intersect. When $c < 0$, the subject exhibits a bias toward reporting targets, and when $c > 0$, a bias toward nontargets.

Sensitivity was indexed as *d'*,

$$d' = \Phi^{-1}(\textit{hit rate}) - \Phi^{-1}(\textit{false alarm rate}).$$

In the signal detection model, *d'* is the difference between the means of the target and nontarget distributions divided by the root mean square of their standard deviations. Because we assume the standard deviation to be equal between the nontarget and target distributions and equal to 1, *d'* is simply the horizontal offset between the means of two standard normal distributions. A larger *d'* indicates better sensitivity. The index *d'* characteristically ranges from zero to infinity, though a negative *d'* value can result from sampling errors.

The results here generalize for other indices of criterion (also known as response bias) used in signal detection theory, such the *likelihood ratio* (β). The index *criterion location* (*c*) has the advantages that *c* is orthogonal to *d'*, well-defined for $d' = 0$, and measured in the same unit as that of *d'* (z-scores) to facilitate comparison (Macmillan and Creelman, 1990).

## Attention task

Two rhesus monkeys (C and Y) performed a variant of the Posner attention task that controls the subject's criterion and sensitivity (Figure 2A). These animals were different from those used in a similar study of V4 (Luo and Maunsell 2015). The subject began each trial by fixating for 400-600 ms within a 2° × 2° square window on a video display (57 cm away, 100 Hz frame rate). Two sample stimuli (full contrast Gabors with $\sigma$ = 1.5° and spatial frequency = 0.7 cycle/°) appeared for 400 ms. Gabors were in opposite hemifields, diametrically opposed across the fixation point, and each 11° from the fixation point. Positions varied from 30° above, 30° below, or directly on the horizontal meridian. After a delay of 150-250 ms, a single test stimulus appeared at a randomly chosen one of the two stimulus locations for 400 ms. The test was equally likely to be the same as the sample that appeared at the same location or different in orientation. If the test differed from the sample, the monkey had to saccade to it within 150-500 ms to receive a juice reward. If the test were the same as the sample, the monkey had to wait to saccade to a second test stimulus that appeared at the same location. The second test always differed from the sample, and it was used to ensure that the monkey was engaged during correct rejection trials. The monkey rarely failed to respond to the second test (< 1%), and trials with such failures were excluded from analyses.

The size of the orientation change was same for all trials in each session and ranged from 45° to 90° across sessions. The sample and the test stimuli in each session could take on only two orientations. The orientation assigned to each sample was randomized on each trial so that the samples had to be inspected to achieve good performance. Stimulus orientation was independent between the two locations.

Each trial was categorized as a hit (H), miss (M), false alarm (FA), or correct rejection (CR) based on animal's response to the first test. A target trial was a H if the monkey responded to the changed test and an M otherwise. A nontarget trial was a FA if the monkey incorrectly responded to the unchanged first test, and it was a CR if the monkey waited to respond to the changed second test. A hit rate ($H$ / [$H$ + $M$]) and a false alarm rate ($FA$ / [$FA$ + $CR$]) were computed for each stimulus location and for each attention condition. Trials with a break in fixation ended immediately and were excluded from analyses.

To control the subject's criterion and sensitivity at each stimulus location, we titrated the reward given for a H and a CR separately at each location. The subject's criterion was primarily influenced by the ratio of the reward for a H to the reward for a CR ("H:CR reward ratio"; Figures 2B and 2C). A larger H:CR reward ratio encouraged a lower criterion, while a lower ratio encouraged a higher criterion. Except for the *low c* condition, the ratio was typically < 1 because subjects preferred responding to the first test over waiting for the second test. The subject's sensitivity at each location was primarily influenced by the average reward size (across H and CR) at that location. A larger average reward at a location resulted in a higher sensitivity at that location. Each day, reward contingencies were varied between two task conditions to isolate a behavioral change in either the subject's criterion or its sensitivity.

The two task conditions of each session were alternated in blocks of 180 trials. In a $\Delta c$ isolation session, the blocks in which the animal was performing with a lower (more negative) $c$ at the stimulus location contralateral to the electrode array are referred to as the "*low c contra*" task condition, and the other blocks the "*high c* contra." Similarly, in a $\Delta d'$ isolation session, blocks in which the subject was performing with a higher $d'$ at the contralateral location was referred to as the "*high d' contra*" task condition, and the other blocks "*low d' contra*" task condition. To encourage stable performance within each block, 50-100 priming trials at the beginning of each block cued the animal to the reward contingencies of that block. Priming trials probed the same stimulus condition 10-20 trials in a row rather than probing the two locations randomly. These trials were excluded from analysis.

After completing training, each monkey was implanted with a pair of 6x8 microelectrode arrays (Blackrock Microsystems) in its right cerebral hemisphere (Figure S2). Behavioral isolation in each recording session was considered satisfactory if the targeted behavioral change was at least four times the off-target behavioral change (median = 26:1). More priming trials were allocated toward stabilizing behavior at the contralateral location than at the ipsilateral location, and as a result, isolation was more consistent at the contralateral location. Behavioral isolation at the contralateral location was achieved in 48 sessions, and in 33 of those sessions, isolation was also considered satisfactory at the ipsilateral location (Figure S1D). Modulations of contra-selective visual neurons were similar between sessions in which the behavior for both stimuli were successfully controlled and sessions in which only the behavior at the contralateral stimulus was successfully controlled.

The 95% confidence intervals of $c$ and $d'$ were computed through bootstrapping. In each of $10^4$ iterations, a random number of hits ($H^{rand}$) was drawn from a binomial distribution based on the observed number of H and M, and a random number of false alarms was similarly drawn.

$$H^{rand} \sim B\left(H+M, \frac{H}{H+M}\right)$$

$$FA^{rand} \sim B\left(FA+CR, \frac{FA}{FA+CR}\right).$$

A hit rate, computed by dividing $H^{rand}$ by the sum of observed H and M, and false alarm rate, computed by dividing $FA^{rand}$ by the sum of observed FA and CR, were used to compute a $c$ and $d'$ for that iteration. The 2.5 and 97.5 percentile across iterations made up the confidence interval.

### Memory-guided saccade task

Each recording session began with the animal performing a memory-guided saccade task. After fixating for 400-600 ms, a saccade target (0.4° white square) appeared for 400 ms in either the left or right hemifield at the center of the Gabors that were subsequently presented in the attention task. After a delay of 750-1000 ms, the fixation point disappeared to cue the animal to make a saccade to within 5° of the remembered location of the target. At least 20 correct trials were completed at each stimulus location in each session.

### Tracking neurons across days

Electrophysiological signals were acquired at 200kHz. Spikes were sorted using the first and second principal components in OfflineSorter (Plexon). Fully isolated clusters were considered to be single units, but partially isolated units and small multiunit clusters were also sorted. All units were included in the tracking analysis regardless of isolation quality.

To track units across sessions, we used the algorithm by Fraser and Schwartz (2012) and the associated software package on MATLAB Central: "Tracking neurons over multiple days," identification no. 30113. Four neurophysiological features—pairwise cross-correlograms, auto-correlogram, waveform shape, and mean firing rate—were used to compute similarity scores for each pair of units between two consecutive sessions. These neuronal metrics were computed using default parameters.

A linear correlation was computed between the average waveforms of the two units being evaluated. The maximum correlation across time shifts was Z-transformed and then taken as the similarity score. Auto-correlograms were computed from 0-100 ms in 5-ms bins, and the similarity score was the Z-transformed linear correlation between two units' auto-correlogram. The similarity score for mean firing rate was the difference between the log of mean rates.

Cross-correlograms were computed from −500 to 500 ms in 100 ms bins. A cross-correlogram was calculated between each of the two units being evaluated and each of all other units present in both sessions. This results in one pair of cross-correlograms for each unit that was present in both sessions (besides the two units being evaluated). A linear correlation was computed between each pair of cross-correlogram and then Z-transformed. The mean z-score was the similarity score. Across the four neuronal metrics, the cross-correlogram provided the most useful information because many neurons were recorded simultaneously in each session, resulting in many cross-correlograms.

A classifier was trained to fit multivariate Gaussian densities to the four similarity scores using a partially supervised expectation-maximization procedure. The decision boundary of the classifier was titrated to achieve similar rates of decoy errors and drop errors (Figures S3A and S3B). A decoy error occurs when one neuron becomes undetectable and another appears, and they are classified as the same unit. A drop error occurs when the same neuron recorded across two consecutive sessions is misclassified as two different units. The decoy error rate was measured using pairs of units from separate microelectrodes, which were most likely different units because the spacing between electrodes was at least 400 μm. The decision boundary was set using the decoy error rate. The drop error was assessed by splitting each recording session into thirds and comparing the first and last thirds as though they were separate sessions. This was a lower bound estimate of the between-session drop error rate. Across a range of decoy error rate, we computed the corresponding drop error rate. Then, we selected the decoy error rate that was closest to the median drop error rate for the tracking procedure.

When multiple units were recorded from the same electrode, there might be multiple possible assignments of identity. When this occurred, an iterative procedure relabeled these units until the summed similarity scores converged to a maximum.

At least two sources of errors were not measured. First, a switch error occurs when the labels are switched between two units, but at least one unit is present on both days. For a switch error to occur, both a decoy error and a drop error would have to occur. Because decoy and drop error rates are close to 1%, the switch error rate is about two orders of magnitude lower than the other two error rates. The second unmeasured error arises when a neuron becomes undetectable for a session but reemerges later. Such a neuron would be given two labels. This error was not corrected because cross-correlograms, the most reliable metric of similarity, become less useful when there are few identical neurons between two sessions. In addition, tracking across non-consecutive sessions greatly increases the chance of decoy errors.

### Spatial selectivity index (SSI)

Spatial selectivity indices (SSIs) were computed using correct trials in the memory-guided saccade task (at least 20 for each location in each session). $SSI_{visual}$ was computed based on a neuron's peak activity during visual target presentation. The 100 ms time window with the maximal firing rate was selected from among three hundred 100 ms windows ranging from [0, 100] ms to [300, 400] ms after the onset of the visual target. The time window of peak response was independently selected for contralateral target trials and ipsilateral target trials. The peak responses to the target at the two positions were compared using a receiver operating characteristic (ROC) analysis. The resulting area under the ROC curve was taken as the SSI.

Other SSIs were computed similarly but using different time windows. $SSI_{delay}$ was computed based the peak firing rate in a 100 ms time window selected from [250, 350] ms to [550, 650] ms relative to target offset. $SSI_{presac}$ was based on a window between [-300, −200] ms to [-100, 0] ms relative to saccade onset, which was defined as the time when the eyes left the fixation window. Finally, $SSI_{postsac}$ was based on a window between [0, 100] ms to [200, 300] ms relative to saccade onset.

The SSI ranges from 0 to 1. SSI > 0.5 indicates a stronger response to the contralateral target or saccade than to the ipsilateral event and SSI < 0.5 indicates the reverse. The time window for indexing selectivity was varied for each neuron according to its peak response because neurons showed a range of visual and saccade-related latencies. This did not bias the SSI away from 0.5

because separate time windows were independently selected for contralateral and ipsilateral events. Varying the counting window for each neuron to select its maximal response has been used previously to characterize spatial selectivity of neuronal responses in prefrontal cortex (Bruce and Goldberg, 1985; Funahashi et al., 1991), though other studies used the same counting window for all neurons (Thompson et al., 2005; Gregoriou et al., 2012).

### Modulation index (MI) and peri-event time histogram (PETH)

A firing rate modulation index (MI) was computed for each neuron during the period 80-480 ms after the onset of the sample stimulus. The MI was computed using trial-averaged firing rates ($\overline{FR}$) from the pair of task conditions in each isolation session:

$$MI_{\Delta c} = \frac{\left(\overline{FR}_{low\ c\ contra} - \overline{FR}_{high\ c\ contra}\right)}{\left(\overline{FR}_{low\ c\ contra} + \overline{FR}_{high\ c\ contra}\right)}$$

$$MI_{\Delta d'} = \frac{\left(\overline{FR}_{high\ d'\ contra} - \overline{FR}_{low\ d'\ contra}\right)}{\left(\overline{FR}_{high\ d'\ contra} + \overline{FR}_{low\ d'\ contra}\right)}.$$

$MI_{\Delta c} > 0$ indicates stronger firing rates when the animal responded more frequently to the contralateral location and less frequently to the ipsilateral location. $MI_{\Delta d'} > 0$ indicates stronger firing rates when the animal was more accurate in discriminating between orientation changes and matches at the contralateral location and less accurate at the ipsilateral location. Modulation indices were computed separately for each of two orientations in each session and then averaged.

To compute the peri-event time histogram for each neuron, spikes were convolved with the trial-averaged histogram with a Gaussian kernel with $\sigma$ = 8 ms. The PETH and modulation index for the sample period were computed using all complete trials (H, M, CR, FA). The PETH aligned to the test includes only trials in which the animal withheld from a response (CR and M), and the PETH aligned to the saccade included only trials in which the animal made a response (H and FA). PETHs and modulation indices were almost identical if only correct trials were used because they far outnumbered incorrect trials. We included error trials in neuronal measurements because our behavioral measurements required error trials. PETH for the test or saccade periods used trials in which the test or the saccade appeared in the neuron's preferred stimulus location, as determined by its $SSI_{visual}$ for the test period and $SSI_{presac}$ for the saccade period.

In Figure 4, the MIs and PETHs of each unit that was recorded during multiple sessions were averaged across sessions of the same isolation.

### Partial correlation between MI and SSI

For the analyses in Figure 5 and Table S1, we included only neurons recorded from the 33/48 sessions in which we achieved behavioral isolation at both the ipsilateral and contralateral stimulus locations. For each unit that was recorded during multiple sessions, an SSI of each type ($SSI_{visual}$, $SSI_{delay}$, $SSI_{presac}$, or $SSI_{postsac}$) and a MI of each type ($MI_{\Delta c}$ or $MI_{\Delta d'}$) was computed for that unit by averaging across sessions. The partial correlation coefficient between a MI and each SSI was computed by first fitting a linear model to explain the MI using the three SSIs being controlled and an intercept. A second model was fitted to explain the SSI that is being correlated using the SSIs being controlled and an intercept. The residuals from the two models were correlated to give the partial correlation.

The 95% confidence interval (*CI*) of each correlation coefficient was computed using the Fisher Z-transformation:

$$CI = [tanh(z - 1.96\sigma),\ tanh(z + 1.96\sigma)]$$

$$z = 0.5log\left(\frac{1+r}{1-r}\right) = \text{atanh}(r)$$

$$\sigma = 1\Big/\sqrt{N-6}$$

where *tanh* is the hyperbolic tangent function, *r* is the correlation coefficient, *N* is the sample size, and *atanh* the inverse hyperbolic tangent function.

### Pairwise correlation and Fano factor

Pairwise correlation was the linear correlation between the spike counts of pairs of simultaneously recorded neurons across trials. Fano factor was computed as each neuron's spike count variance across trials divided by its spike count average across trials. Both

pairwise correlation and Fano factor were based on spike counts during the period 80-480 ms after the onset of the sample stimulus during all complete trials (H, M, FA, and CR). Both were computed separately for each of the two sample orientations and then averaged. For each unit that was recorded during multiple sessions, Fano factor values were averaged across sessions of the same isolation. Similarly, for each pair of unit that was present in multiple sessions, correlation values were also averaged across sessions of the same isolation.

Because pairwise correlations can be negative, the standard modulation index was not used to index changes in correlations. Instead, differences in pairwise correlation between task conditions were normalized to the average correlation across pairs in either the *high c contra* condition (for pairs from *Δc* isolations) or the *low d' contra* condition (for pairs from *Δd'* isolations). Normalized differences were also computed for Fano factors, even though Fano factors are always positive, so that the size of Fano factor modulations can be compared to that of pairwise correlation modulations.

### Comparison between LPFC and V4

In Figure 7, we compared the activity of contra-selective visual neurons in LPFC to the activity of V4 neurons that were recorded in the experiments described in Luo and Maunsell (2015). In the V4 dataset, the sample stimuli appeared for only 200 ms rather than 400 ms in the LPFC dataset and no memory-guided saccade task was performed. To facilitate comparison, firing rates of contra-selective visual neurons in LPFC were computed between 80 ms to 280 ms after sample onset. Moreover, we computed an approximate $SSI_{visual}$ for each V4 unit. But because V4 neurons have entirely contralateral receptive fields, $SSI_{visual}$ was approximated by performing a ROC analysis comparing firing rates to the sample stimulus (60 ms to 260 ms after sample onset) to the baseline response during the 200 ms before sample onset.

The attention-related change in firing rate or Fano factor of each neuron was normalized to the average across all neurons in that area in either the *high c contra* or *low d' contra* task condition. Similarly, the change in the correlation of each neuronal pair was normalized to the average correlation across all pairs in that area in either the *high c contra* or *low d' contra* task condition.

Two models were fitted to the normalized changes in firing rates in LPFC and V4 (Figures 7B and 7C). The distribution of normalized firing rate changes for each brain area and for each attention component was modeled as a normal distribution. In the first model, the means of *Δd'*-related changes in LPFC, *Δc*-related changes in LPFC, and *Δd'*-related changes in V4 were allowed to vary. However, the mean of *Δc*-related changes in V4 was constrained to be zero. The standard deviation ($\sigma$) of each distribution was constrained to be the sample standard deviation of each distribution. If $\sigma$ were allowed to vary, the resulting log-likelihood ratios would be little changed and the fitted $\sigma$ would be similar to the sample standard deviation.

$$\Delta FR'_{\Delta d',\,LPFC} \sim \mathcal{N}(\beta_1, \sigma_{sample})$$

$$\Delta FR'_{\Delta c,LPFC} \sim \mathcal{N}(\beta_2, \sigma_{sample})$$

$$\Delta FR'_{\Delta d',V4} \sim \mathcal{N}(\beta_3, \sigma_{sample})$$

$$\Delta FR'_{\Delta c,V4} \sim \mathcal{N}(0, \sigma_{sample}).$$

In the second model, the means of *Δd'*-related changes and *Δc*-related changes in LPFC are again allowed to vary. However, the means of changes in V4 are constrained to be related to the changes in LPFC by a multiplicative factor, which was allowed to vary.

$$\Delta FR'_{\Delta d',LPFC} \sim \mathcal{N}(\beta_1, \sigma_{sample})$$

$$\Delta FR'_{\Delta c,LPFC} \sim \mathcal{N}(\beta_2, \sigma_{sample})$$

$$\Delta FR'_{\Delta d',V4} \sim \mathcal{N}(\beta_3 * \beta_1, \sigma_{sample})$$

$$\Delta FR'_{\Delta c,V4} \sim \mathcal{N}(\beta_3 * \beta_2, \sigma_{sample}).$$

Each model had three parameters, and they were fitted using maximum likelihood estimation and using the *fminunc* function in MATLAB. When fitting the models, because different numbers of neurons were recorded in different brain areas and in different types of behavioral isolation, we subsampled the neurons such that each combination of brain area and isolation type had the same number of neurons, which was the minimum across combinations. For example, for $SSI_{visual} > 0.75$, the minimum was 99 neurons recoded in LPFC during *Δc* isolations. Subsampling followed by model fitting were repeated for $10^4$ iterations, and we reported the median, 2.5

percentile, and 97.5 percentile of the bootstrapped distribution of likelihood ratios. These same analyses were performed for normalized changes in pairwise correlation and in Fano factor.

To determine whether individual units in V4 and LPFC were significantly modulated in correlation with either $\Delta c$ or $\Delta d'$ (Figures 7D and 7E), we fitted a Poisson model to the neuron's trial-by-trial spike count during the sample period. We did not simply perform a statistical test comparing the neuron's firing rates across trials in the two different task conditions because many V4 units showed slow change in firing rates (either increase or decrease) over the course of a recording session. The probability of the number of spikes emitted by each neuron during the sample period in each trial was modeled as:

$$P(K) = \lambda^K e^{-\lambda} / K!$$

$$\lambda = \beta_0 + \beta_1\, t + \beta_2\, \theta + \beta_3\, attention$$

where $K$ is the spike count, $t$ is the trial number, $\theta$ is either a 0 or 1 indexing the orientation of the sample stimulus, and *attention* = 1 for either the *low c contra* condition or *high d' contra* condition, and 0 otherwise. An additional restricted model was fitted to the trial-by-trial spike counts, with the constraint $\beta_3 = 0$. The restricted model and the unrestricted model were fitted using maximum likelihood estimation, and the two models were compared using a likelihood ratio test, thereby giving the p value of a neuron being significantly modulated.

### Correlation between Δc- and Δd'-related firing rate MI

Firing rates modulation indices (MI) in Figure 8 were computed by counting spikes 80-280 ms after sample onset for units recorded in LPFC and 60-260 ms after sample onset for units in V4. For each unit that was recorded for multiple sessions, its MI from different sessions were paired. These pairs of MI were pooled across units and sorted according to whether a pair came from the same behavioral isolation (same-isolation MI pair) or different behavioral isolations ($\Delta c/\Delta d'$ MI pair).

Different number of sessions intervened between the earlier and later session of each MI pair (0-23 for LPFC; 0-43 for V4). On average, more sessions intervened between $\Delta c/\Delta d'$ MI pairs than between same-isolation MI pairs because the same behavioral isolation occurred at least five sessions in a row (Figures S7A and S7B). To control for this difference, we matched the distributions of intervening sessions between the two types of MI pairs (Churchland et al., 2010). For each value of intervening session, the MI pairs of the type that had more pairs were subsampled to have the same number of pairs as the other type. Subsampled pairs across all values of intervening session were pooled to compute a correlation coefficient. Random subsampling was repeated 1000 times. The median correlation coefficient across the 1000 iterations and its associated confidence interval (computed using the Fisher Z-transform) were reported in Figure 8C.

The normalized $\Delta c/\Delta d'$ correlation was obtained by dividing the $\Delta c/\Delta d'$ correlation by the same-isolation correlation, and its 95% confidence interval was computed through bootstrapping. First, correlation coefficients ($\rho$) were Z-transformed:

$$z_{\Delta c/\Delta d'} = \operatorname{atanh}\left(\rho_{\Delta c/\Delta d'}\right)$$

$$z_{same-iso} = \operatorname{atanh}\left(\rho_{same-iso}\right).$$

Then, the standard error of the Z-transformed coefficient was computed as:

$$\sigma = 1 \Big/ \sqrt{N-3}.$$

Because both types of MI pairs had the same sample size after distribution-matching, the two Z-transformed coefficients had the same standard error. Then, $10^5$ random samples were drawn from each of these normal distributions:

$$X_{c/\Delta d'} \sim N\left(z_{\Delta c/\Delta d'}, \sigma\right)$$

$$X_{same-iso} \sim N\left(z_{same-iso}, \sigma\right).$$

These random samples were transformed into correlation coefficients, and a ratio was taken between each pair of random samples:

$$r_{c/\Delta d'} \sim \tanh\left(X_{c/\Delta d'}\right)$$

$$r_{same-iso} \sim \tanh\left(X_{same-iso}\right)$$

$$R = \frac{r_{c/\Delta d'}}{r_{same-iso}}.$$

The 2.5 and 97.5 percentiles of the resulting distribution of ratios was taken as the 95% confidence interval.

### Linear Fisher information

In Figure S5, linear Fisher information was computed following the methods in Kanitscheider et al. (2015). In each session, we computed the differences in the trial-averaged firing rates to the Gabor stimuli of two different orientations for all units with $SSI_{visual} > 0.75$.

$$d\mu \equiv \overline{FR}(\theta^+) - \overline{FR}(\theta^-)$$

$d\mu$ is a vector with a number of elements equal to the number of units with $SSI_{visual} > 0.75$ in that session. We further computed the inverse of the average of the covariance matrices for each orientation.

$$S^{-1} \equiv \left(\frac{S^+ + S^-}{2}\right)^{-1}.$$

Linear Fisher information was computed as

$$I = \frac{d\mu^T}{d\theta} S^{-1} \frac{d\mu}{d\theta} \left(\frac{2T - N - 3}{2T - 2}\right) - \frac{2N}{Td\theta^2}$$

where $d\theta$ is the difference in orientation between the two stimuli in radians, $T$ is the number of trials, and $N$ is the number of units. For sessions in which the number of trials differed for the two orientations, we subsampled the condition with more trials across 1000 iterations and averaged the estimated linear Fisher information across iterations.

Shuffled linear Fisher information was computed as

$$I_{shuffle} = \sum_i \frac{(d\mu_i/d\theta)^2}{s_i^2} \left(\frac{T - 2}{T - 1}\right) - \frac{2N}{Td\theta^2}$$

where $d\mu_i$ is the trial-averaged firing rate difference between the two stimuli for neuron $i$ and $s_i^2$ is its sample variance. Results were similar across a broad range of minimum $SSI_{visual}$ (not shown).

## QUANTIFICATION AND STATISTICAL ANALYSES

The details of all statistical tests performed are described in the Method Details, in the main text, or in the figure captions. Details include the number of neuronal units and recording sessions used for each analysis. Error bars and shaded regions on plots indicate either mean ± SEM (Figures 3, 4, 6, and 7A) or 95% confidence interval (Figures 2, 5, 7C, and 8).

## DATA AND SOFTWARE AVAILABILITY

The raw data are available at Mendeley Data, https://doi.org/10.17632/7ns5zz87bp.1.

**Neuron, Volume *97***

## Supplemental Information

# Attentional Changes in Either Criterion or Sensitivity Are Associated with Robust Modulations in Lateral Prefrontal Cortex

**Thomas Zhihao Luo and John H.R. Maunsell**

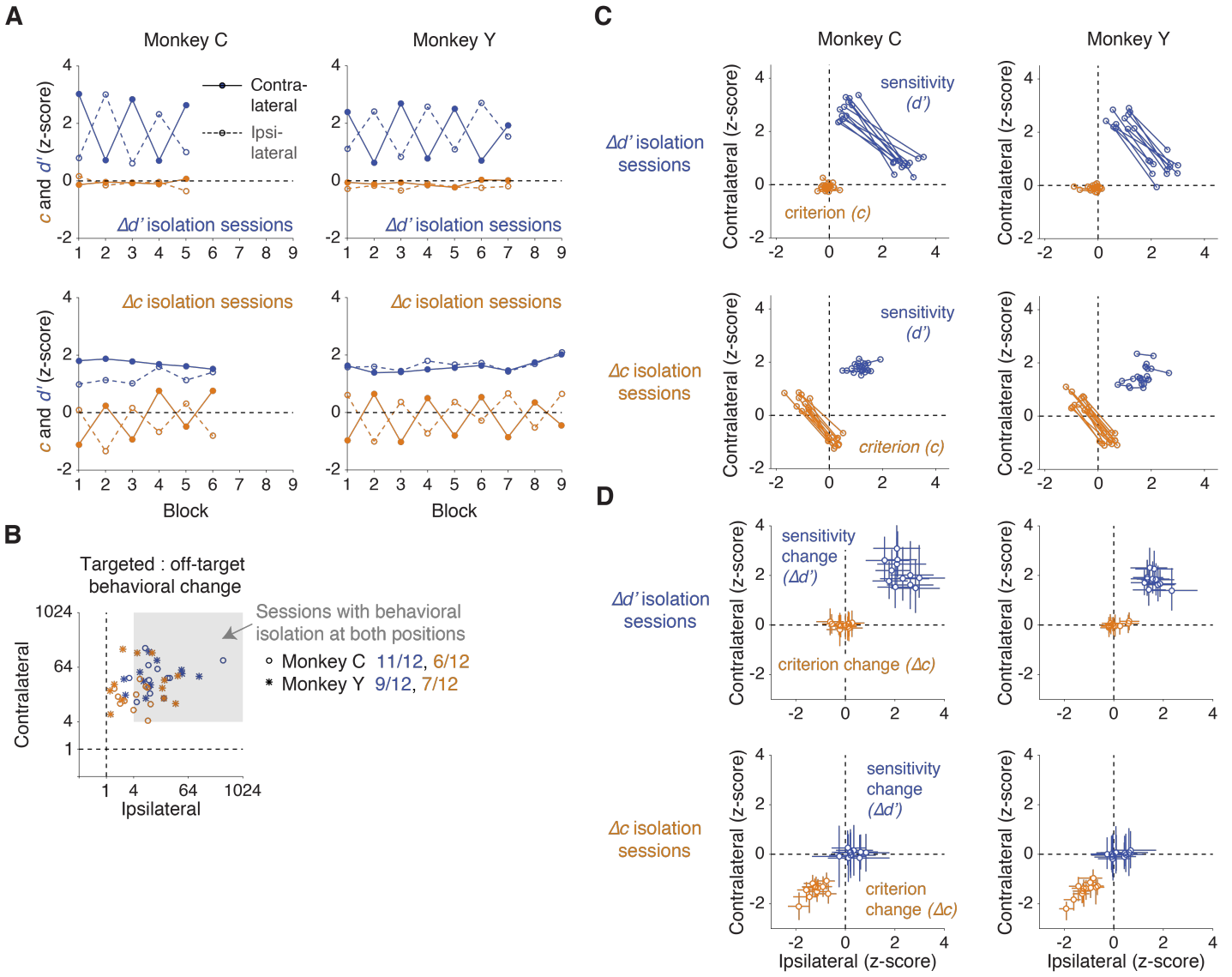


**Figure S1. Behavioral performance at the stimulus positions contralateral and ipsilateral to the implanted electrode arrays. Related to Figure 2.** (**A**) During Δ*d'* isolation sessions, the subject's *d'* at the contralateral and ipsilateral locations alternated across blocks between being either high or low while the subject's *c* at both locations remained stable. Conversely, during *Δc* isolations, the subject's *c* at the contralateral and ipsilateral locations changed in counterphase while the subject's *d'* at both locations remained stable. The counterphase relationship between the two locations indicated that the isolated changes were spatially selective. Data represent average behavior across sessions. Half of the sessions started with the first block attending to the location contralateral to the implanted array, and the other half started with attending to the ipsilateral location. Sessions that started with attending-ipsi were shifted by one block in this plot to align with sessions that started with attending-contra. (**B**) The ratio of targeted behavioral change to off-target change. In a Δ*d'* isolation session, the target:off-target ratio was Δ*d'* / Δ*c*, and in a Δ*c* isolation session, this ratio was Δ*c* / Δ*d'*. A ratio > 4 was considered to indicate a successful isolation. The median ratio was 26:1 at the contralateral location and 8:1 at the ipsilateral location. During 2/48 sessions, not enough errors were obtained at the ipsilateral location to compute *c* and *d'*, and isolation was considered unsuccessful at the ipsilateral location. (**C**) The subject's criterion (*c*) and sensitivity (*d'*) in each session. Each marker indicates either *c* (orange) or *d'* (blue) in one task condition of each session. Markers representing the same session are connected by a line. Error bars were omitted for clarity. (**D**) The subject's criterion change (*Δc*) and sensitivity change (*Δd'*) in each session.

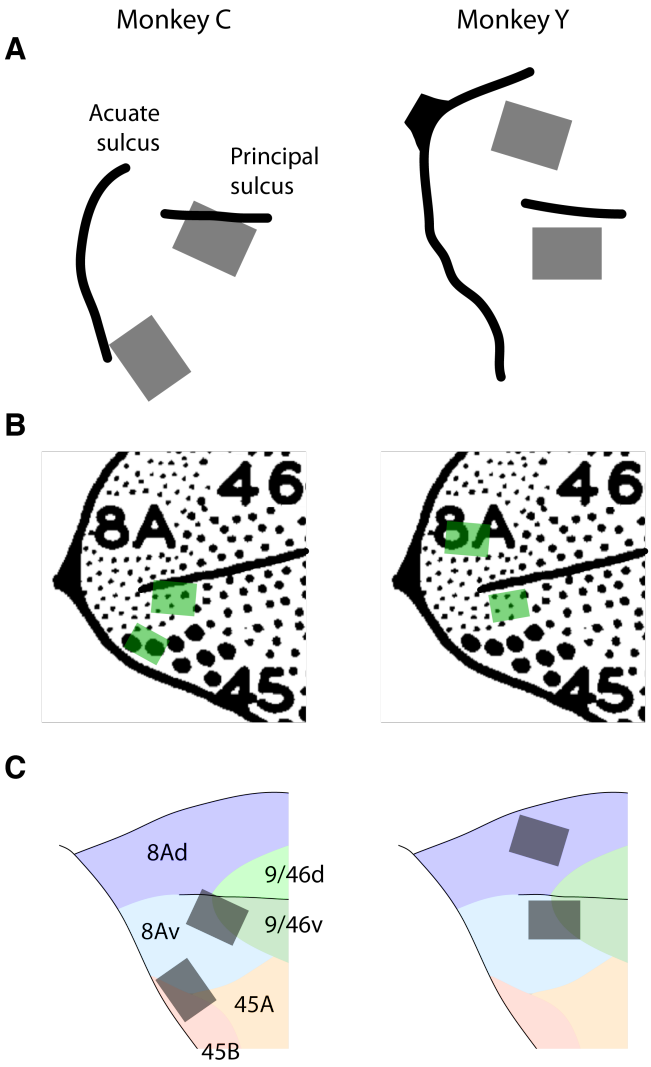


**Figure S2. Anatomical locations of microelectrode arrays. Related to STAR Methods.** (**A**) Microelectrode arrays placement relative to frontal sulci. Diagrams were traced from photographs taken during surgery. (**B**) Diagrams of the arrays superimposed approximately onto the map from (Walker, 1940). For monkey C, the dorsal array extends into areas 8A and 46, and the ventral array extends into areas 8A and 45. For monkey Y, the dorsal array is within area 8A and the ventral array extends into area 8A and 46. (**C**) Arrays superimposed approximately onto the map from (Petrides and Pandya, 1999). For monkey C, the dorsal array extends into areas 8A and 9/46, and the ventral array extends into areas 8A, 45A, and 45B. For monkey Y, the dorsal array is within area 8A and the ventral array extends into area 8A and 9/46. We refer to these areas as "lateral" prefrontal cortex rather than "dorsolateral" prefrontal cortex because area 45 has been described to be part of "ventrolateral" prefrontal cortex (Petrides and Pandya, 2002).

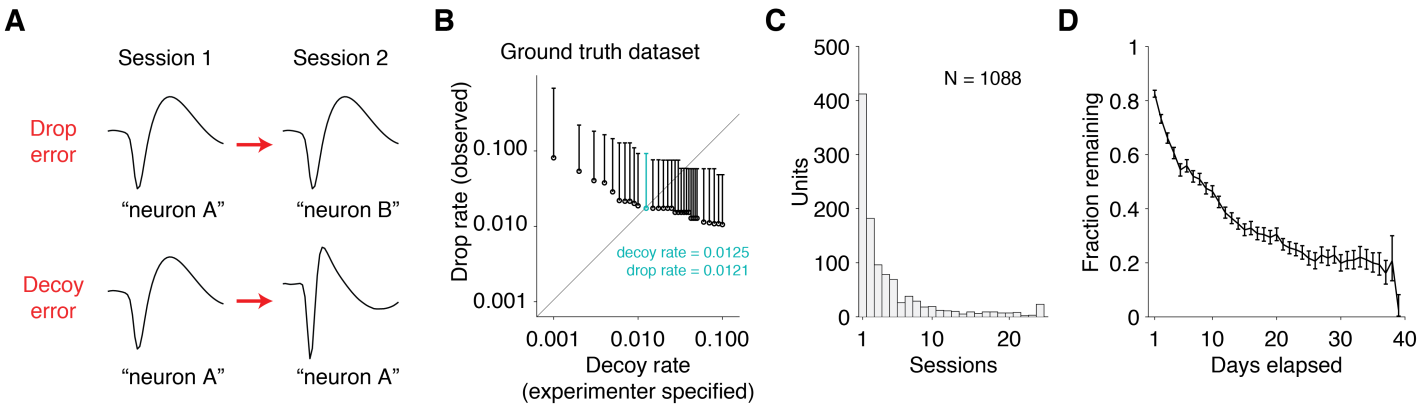


**Figure S3. Tracking neurons across sessions. Related to STAR Methods.** (**A**) Two types of error can occur while trying to track the same neuron over multiple sessions. A drop error occurs when the same neuron recorded during two sessions is classified as two different units. A decoy error occurs when two different neurons are classified as the same unit. (**B**) To assess the accuracy of the tracking algorithm, a ground truth dataset was generated by splitting each recording session into thirds and then comparing the first and last thirds as though they were separate sessions. Each pair of thirds of the same session was processed using different target decoy rates, and the resulting drop rate was measured. The decoy rate of each classifier was specified as the fraction of units from different electrodes (which are almost certainly not the same unit because electrodes were at least 400 μm apart) that would be classified as identical. Each circle indicates the median drop rate across comparisons, and the upper error bar indicates the maximum drop rate across comparisons. The minimum drop rate was 0 for all tested target decoy rates and therefore not plotted. The decoy rate that was most similar to the observed drop rate (cyan) was selected for classifying the full dataset. Note that the drop rate measured here is a lower bound of the drop rate that occurs when tracking units across days. (**C**) The number of consecutive sessions across which a neuron was tracked. Among the 4,843 units encountered across all days, 1,088 (22% of 4,843) were classified to be unique, and the remaining 3,755 were repeated recordings of the 676 unique units (62% of 1,088) that were recorded for more than one session. (**D**) The fraction of neurons in an earlier session that was also recorded in a later session, sorted by the intervening number of days between the two sessions. Error bars indicate 95% binomial confidence interval.

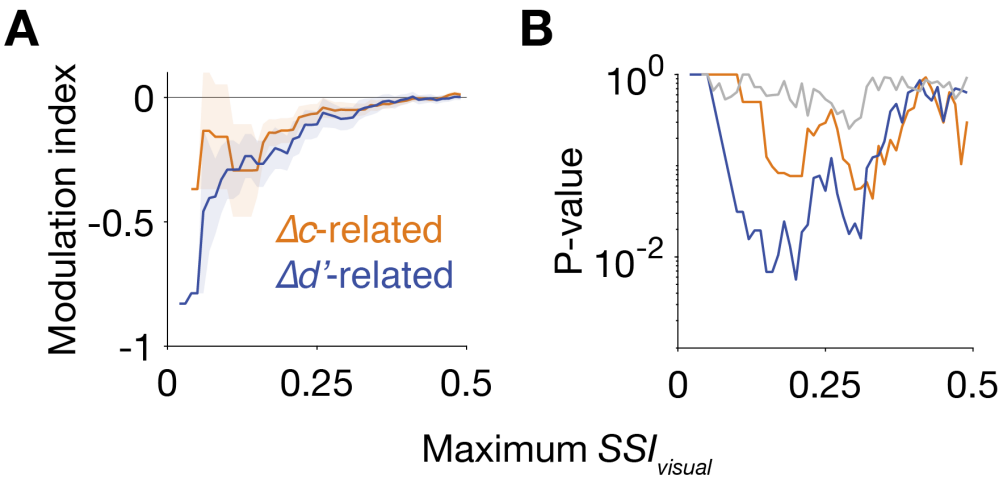


**Figure S4. Modulation indices of ipsi-selective visual neurons. Related Figure 4.** (**A**) Modulation index (MI) of firing rates during the period 80-480 ms after sample onset as function of the maximum $SSI_{visual}$ of the units included for analysis. A $\Delta c$-related modulation index < 0 indicates stronger firing rates when the animal responded more frequently to ipsilateral location. A $\Delta d'$-related MI < 0 indicates stronger firing rates when the animal discriminated more accurately at the ipsilateral position. Shading indicates mean ± SEM across units. Only the 33/48 sessions in which behavior was isolated at both the ipsilateral and contralateral locations were included for this analysis. The sample sizes for maximum $SSI_{visual}$ = 0.49 were N = 278 ($\Delta c$) and N = 349 ($\Delta d'$). (**B**) P-values were computed using the Wilcoxon signed-rank test for a zero median in the distribution of $\Delta c$-related modulation indices (orange) or $\Delta d'$-related indices (blue) and from the Wilcoxon rank-sum test of $\Delta c$- and $\Delta d'$-related indices having the same median (gray).

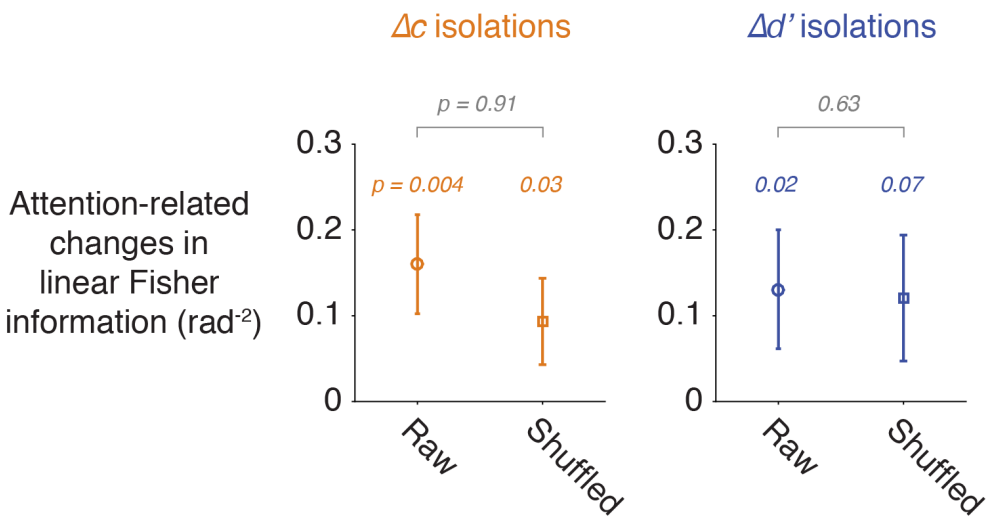


**Figure S5. Linear Fisher information, a measure of discriminability of the stimuli by the neural population, increased modestly in association with either decreases in criterion or increases in sensitivity. Related Figure 6.** Linear Fisher information is defined as the inverse of the variance of the locally optimal linear decoder of the neural population responses. Higher linear Fisher information indicates better discrimination of the stimuli. Shuffled information quantifies information encoding in the same neural population after correlations between neurons have been removed. Error bars indicate mean ± SEM across 24 sessions.

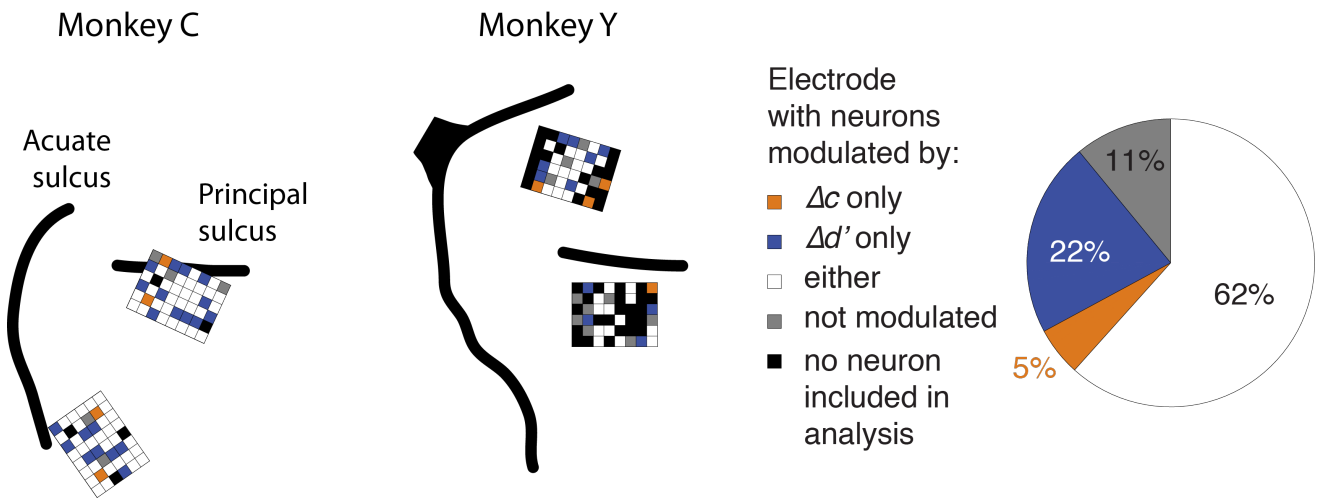


**Figure S6. Both *Δc*-related and *Δd'*-related modulations were observed in all microelectrode arrays and could be observed in the same electrodes (across different days). Related to Figure 8.** Each colored square indicates whether the units recorded from an electrode were modulated when the animal changed only its criterion (orange), only its sensitivity (blue), or either (white). Gray indicates that units were not modulated in association with either behavioral change, and black indicates that no unit from that electrode was included for this analysis. A neuron was included for this analysis if its firing rates were modulated during the sample period of the attention task relative to the fixation period at $p = 0.01$ (Wilcoxon rank-sum test) and also have a trial-averaged firing rate of at least 5 Hz during any task period. Sixty-five percent of units recorded was included. Neurons were then sorted according to their electrode and the type of isolation session when they were recorded. For each group of neurons from the same electrode and same isolation, an ANOVA with factors *neuron* and *attention* was performed. The *attention* factor indicates whether firing rate was modulated in relation to either *Δc* or *Δd'*. An electrode was considered to be modulated if the main effect of *attention* had a p-value less than 0.01. The pie chart includes only significantly modulated electrodes.

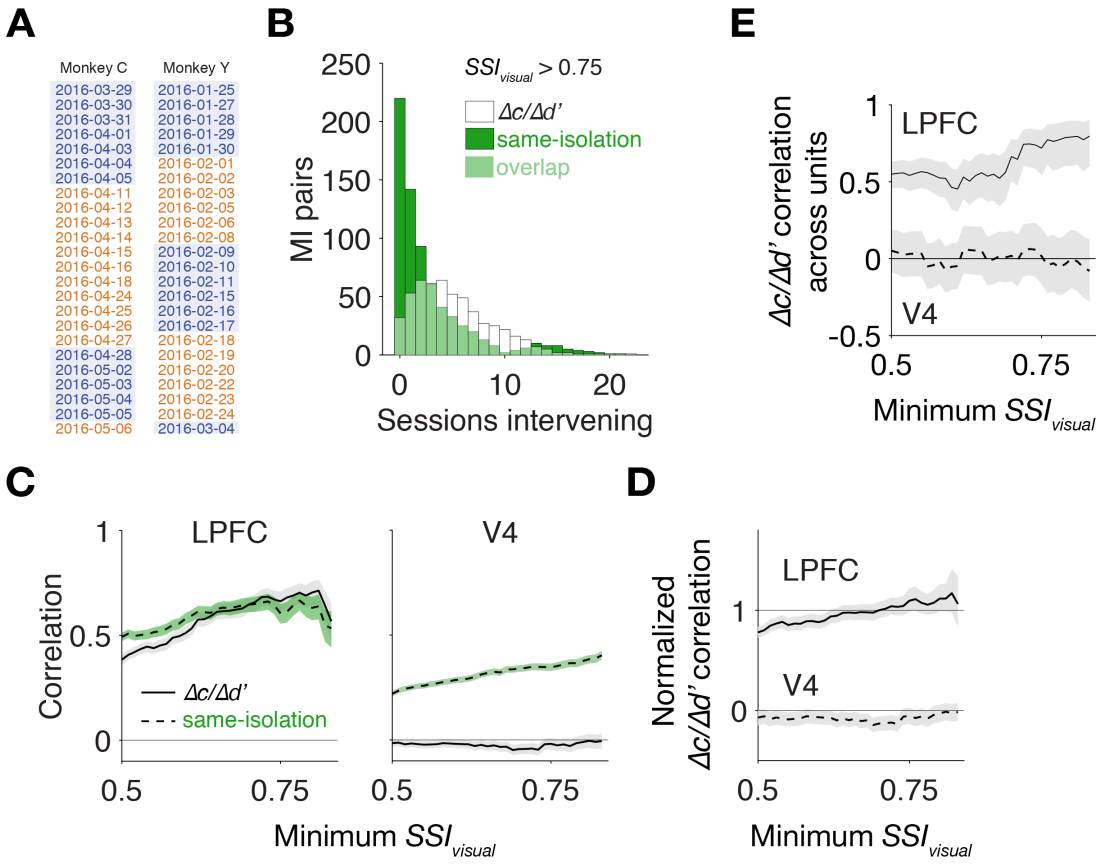


**Figure S7. Additional analyses related to Figure 8.** (**A**) The dates of the 24 *Δc*-isolations (orange) and *Δd'*-isolations (blue) during which neurons from LPFC were recorded. Because the type of isolation switched only after at least five successful sessions, *Δc*/*Δd'* sessions pairs were on average spaced further apart in time. (**B**) The number of intervening sessions between each pair of MI. Each MI pair was recorded from the same unit, and all pairs of MI's were pooled across all units. To match the distributions, for each bin, a number of MI pairs was randomly selected from each of the *Δc*/*Δd'* distribution (white) and same-isolation distribution (green), and this number was equal to the overlap between the two distributions (light green) in that bin. Subsets of MI pairs were pooled across bins to calculate the correlation coefficient. Random subsampling was repeated 1,000 times to provide 1,000 correlation coefficients and confidence intervals. The median correlation coefficient and its associated confidence interval were shown in **Figures 8C-D**. (**C-D**) The same analyses in **Figure 8C-D** performed *without* distribution-matching, showing that it did not qualitatively change the results. (**E**) An alternative analysis that computed correlations across units rather than pairs of MIs similarly found a difference in *Δc*/*Δd'* correlations between LPFC and V4. In this analysis, a single *Δc*-related MI and a single *Δd'*-related MI was computed for each unit by averaging across MIs from different sessions in which that unit was recorded. Shading indicates 95% confidence interval.

| Model | Δd'-related modulations | | Δc-related modulations | |
|---|---|---|---|---|
| | Variance explained (%) | BIC | Variance explained (%) | BIC |
| *MI ~ b0* | <1 | 251.48 | <1 | 100.68 |
| *MI ~ b0*visual* | 14.6 | 35.23 | 9 | 18.39 |
| *MI ~ b0*delay* | <1 | 256.23 | <1 | 110.3 |
| *MI ~ b0*presac* | <1 | 256.69 | <1 | 129.05 |
| *MI ~ b0*postsac* | <1 | 321.57 | <1 | 162.22 |
| *MI ~ b0 + b1*visual* | 15.4 | 29.13 | 10.8 | 8.25 |
| *MI ~ b0 + b1*delay* | 3.8 | 206.04 | 4.1 | 71.39 |
| *MI ~ b0*visual + b1*delay* | 15.4 | 28.36 | 10.4 | 11.52 |
| *MI ~ b0*visual + b1*delay + b2*visual:delay* | 16.2 | 23.62 | 11.8 | 4.5 |
| *MI ~ b0 + b1*presac* | 3.8 | 205.11 | 2.4 | 86.65 |
| *MI ~ b0*visual + b1*presac* | 16.3 | 13.91 | 10.6 | 10.19 |
| *MI ~ b0*visual + b1*presac + b2*visual:presac* | 16.9 | 12.1 | 11.7 | 5.59 |
| *MI ~ b0*delay + b1*presac* | 1 | 244.79 | <1 | 113.33 |
| *MI ~ b0*delay + b1*presac + b2*delay:presac* | 4.6 | 200.65 | 4.1 | 78.15 |
| *MI ~ b0 + b1*postsac* | <1 | 251.44 | <1 | 107.42 |
| *MI ~ b0*visual + b1*postsac* | 15.7 | 23.29 | 9.1 | 24.41 |
| *MI ~ b0*visual + b1*postsac + b2*visual:postsac* | 16.5 | 18.41 | 10.7 | 15.71 |
| *MI ~ b0*delay + b1*postsac* | <1 | 250.24 | <1 | 116.59 |
| *MI ~ b0*delay + b1*postsac + b2*delay:postsac* | 4.5 | 202.26 | 4.2 | 76.74 |
| *MI ~ b0*presac + b1*postsac* | <1 | 257.79 | <1 | 135.77 |
| *MI ~ b0*presac + b1*postsac + b2*presac:postsac* | 4.3 | 206.21 | 2.3 | 94.21 |
| *MI ~ b0 + b1*visual + b2*delay* | 16.2 | 23.9 | 11.9 | 3.4 |
| *MI ~ b0 + b1*visual + b2*delay + b3*visual:delay* | 16.2 | 30.8 | 12.1 | 8.91 |
| *MI ~ b0 + b1*visual + b2*presac* | 17 | 10.74 | 11.9 | 3.43 |
| *MI ~ b0 + b1*visual + b2*presac + b3*visual:presac* | 17.3 | 12.23 | 13 | 0 |
| *MI ~ b0 + b1*delay + b2*presac* | 4.8 | 198.35 | 4.3 | 76.29 |
| *MI ~ b0 + b1*delay + b2*presac + b3*delay:presac* | 5 | 202.78 | 4.4 | 81.42 |
| *MI ~ b0*visual + b1*delay + b2*presac* | 16.4 | 20.29 | 10.9 | 13.67 |
| *MI ~ b0*visual + b1*delay + b2*presac + b3*visual:delay + b4*visual:delay:presac* | 17.4 | 17.46 | 12.9 | 7.75 |
| *MI ~ b0*visual + b1*delay + b2*presac + b3*visual:delay + b4*visual:presac* | 17.6 | 15.11 | 12.5 | 11.63 |
| *MI ~ b0 + b1*visual + b2*postsac* | 16.5 | 17.5 | 10.8 | 14.73 |
| *MI ~ b0 + b1*visual + b2*postsac + b3*visual:postsac* | 16.6 | 24.02 | 10.9 | 20.93 |
| *MI ~ b0 + b1*delay + b2*postsac* | 4.7 | 199.8 | 4.1 | 78.12 |
| *MI ~ b0 + b1*delay + b2*postsac + b3*delay:postsac* | 4.9 | 204.7 | 4.3 | 82.99 |
| *MI ~ b0*visual + b1*delay + b2*postsac* | 16.8 | 12.49 | 10.5 | 17.62 |
| *MI ~ b0*visual + b1*delay + b2*postsac + b3*delay:postsac* | 17.5 | 8.7 | 12.1 | 8.67 |
| *MI ~ b0*visual + b1*delay + b2*postsac + b3*visual:delay:postsac* | 17.6 | 7.77 | 12 | 10.1 |
| *MI ~ b0 + b1*presac + b2*postsac* | 4.3 | 205.27 | 2.4 | 93.12 |
| *MI ~ b0 + b1*presac + b2*postsac + b3*presac:postsac* | 4.3 | 212.19 | 2.4 | 99.57 |
| *MI ~ b0*visual + b1*presac + b2*postsac* | 17.4 | 2.97 | 10.6 | 16.92 |
| *MI ~ b0*visual + b1*presac + b2*postsac + b3*presac:postsac* | 18 | 0 | 11.9 | 10.26 |
| *MI ~ b0*delay + b1*presac + b2*postsac* | 1.8 | 241.57 | <1 | 120 |
| *MI ~ b0*delay + b1*presac + b2*postsac + b3*delay:postsac* | 5.4 | 196.65 | 4.4 | 81.47 |
| *MI ~ b0*delay + b1*presac + b2*postsac + b3*presac:postsac* | 5.5 | 195.14 | 4.2 | 83.51 |
| *MI ~ b0 + b1*visual + b2*delay + b3*presac* | 17 | 17.24 | 12.3 | 6.97 |
| *MI ~ b0 + b1*visual + b2*delay + b3*presac + b4*visual:presac* | 17.3 | 18.56 | 13.3 | 3.47 |
| *MI ~ b0 + b1*visual + b2*delay + b3*postsac* | 17.5 | 8.58 | 12 | 9.91 |
| *MI ~ b0 + b1*visual + b2*delay + b3*postsac + b4*delay:postsac* | 17.5 | 15.8 | 12.3 | 13.57 |
| *MI ~ b0 + b1*visual + b2*delay + b3*postsac + b4*visual:delay* | 17.6 | 14.66 | 12.1 | 15.49 |
| *MI ~ b0 + b1*visual + b2*presac + b3*postsac* | 18 | 0.02 | 11.9 | 10.2 |
| *MI ~ b0 + b1*visual + b2*presac + b3*postsac + b4*visual:presac* | 18.3 | 2.67 | 13 | 6.65 |
| *MI ~ b0 + b1*delay + b2*presac + b3*postsac* | 5.6 | 194.63 | 4.3 | 83.06 |
| *MI ~ b0 + b1*delay + b2*presac + b3*postsac + b4*delay:postsac* | 5.7 | 200.3 | 4.5 | 87.41 |
| *MI ~ b0 + b1*delay + b2*presac + b3*postsac + b4*delay:presac:postsac* | 5.7 | 199.73 | 4.3 | 89.78 |
| *MI ~ b0*visual + b1*delay + b2*presac + b3*postsac* | 17.6 | 7.8 | 10.9 | 20.29 |
| *MI ~ b0*visual + b1*delay + b2*presac + b3*postsac + b4*visual:presac + b5*delay:postsac* | 18.3 | 10.26 | 13.5 | 7.74 |
| *MI ~ b0*visual + b1*delay + b2*presac + b3*postsac + b4*visual:delay + b5*visual:presac* | 18.8 | 0.86 | 12.5 | 18.38 |
| *MI ~ b0 + b1*visual + b2*delay + b3*presac + b4*postsac* | 18.2 | 5.08 | 12.3 | 13.71 |
| *MI ~ b0 + b1*visual + b2*delay + b3*presac + b4*postsac + b5*visual:presac* | 18.4 | 7.52 | 13.3 | 10.21 |
| *MI ~ b0 + b1*visual + b2*delay + b3*presac + b4*postsac + b5*visual:delay + b6*visual:presac* | 19 | 5.86 | 13.4 | 15.69 |

**Table S1. Linear models of firing rate modulation indices using SSI's as predictor variables. Related to Figure 5.** Modulation indices of firing rate (*MI*) were computed from spike counts during the sample period in the attention task. The optimal model was selected according to the lowest Bayesian Information Criterion (BIC). BIC values are displayed relative to the minimum across models. To simplify presentation, for each set of models that have the same variables but different combinations of interaction terms, we show only the model that has the lowest BIC for each type of modulation. The number of models considered was 4,257.