Demands of visual processing hierarchy shape laminar compartmentalization of attention modulation of luminance contrast in area V4

Xiang Wang[1], Anirvan S. Nandy[1,3†] and Monika P. Jadi[1,2,3,*,†]
[1]Interdepartmental Neuroscience Program, Yale University, New Haven, CT 06511
[2]Department of Psychiatry, Yale University, New Haven, CT 06511
[3]Department of Neuroscience, Yale University, New Haven, CT 06511
†senior author
*corresponding author (monika.jadi@yale.edu)

ABSTRACT

Contrast is a key feature of the visual scene that aids object recognition. Attention has been shown to selectively enhance the responses to low contrast stimuli in visual area V4, a critical hub that sends projections both up and down the visual hierarchy. Veridical encoding of contrast information is a key computation in early visual areas, while later stages encode higher level features that benefit from improved sensitivity to low contrast. How area V4 meets these distinct information processing demands in the attentive state is not known. We found that attentional modulation of contrast responses in area V4 is cortical layer and cell-class specific. Putative excitatory neurons in the superficial output layers that project to higher areas show enhanced boosting of low contrast information. On the other hand, putative excitatory neurons of deep output layers that project to early visual areas exhibit contrast-independent scaling. Computational modeling revealed that such layer-wise differences may result from variations in spatial integration extent of inhibitory neurons. These findings reveal that the nature of interactions between attention and contrast in V4 is highly compartmentalized, in alignment with the demands of the visual processing hierarchy.

INTRODUCTION

Voluntary attention is essential for sensory guided behavior and memory formation (Petersen and Posner, 2012). Failures in sensory processing and selective attention are aspects of many mental illnesses, including schizophrenia and mood disorders (Fioravanti et al., 2005; McIntyre et al., 2010; Neuchterlein et al., 1991). Visual spatial attention plays a critical role in visual sensory processing: It allows improved perception of behaviorally relevant target stimuli among competing distractors by boosting the apparent visibility of the target (Carrasco et al., 2004). At the neuronal level, attention modulates the activity of cortical neurons that encode an attended visual stimulus at various stages of visual processing (Bisley and Goldberg, 2003; Ghose and Maunsell, 2008; Moran and Desimone, 1985; Motter, 1993; Reynolds et al., 1999; Treue and Martinez Trujillo, 1999; Treue and Maunsell, 1996). In visual areas such as V4 and MT, attention modulates neuronal mean firing rates, increases their firing reliability, and reduces the co-variability among pairs of neurons (Cohen and Maunsell, 2009; Mitchell et al., 2007, 2009; Reynolds and Chelazzi, 2004; Treue and Martinez Trujillo, 1999). However, the computational principles that underlie the activity of neuronal populations that represent both sensory information and the attentional state remain poorly understood (Moore and Zirnsak, 2017; Reynolds and Chelazzi, 2004).

Object recognition is mediated by a hierarchy of cortical visual processing areas that form the ventral visual stream. Contrast is a key feature of the visual scene that aids object recognition, and the encoding of contrast information is one of the most important computations performed by early visual areas. On the other hand, visual features represented in higher areas such as the inferotemporal (IT) cortex benefit from improved sensitivity to low contrast stimuli (Avidan et al., 2002; Rolls and Baylis, 1986). Visual area V4 is a critical hub in the ventral stream that sends feedforward projections to areas such as IT and feedback projections to early visual processing areas (Anderson and Martin, 2006; Douglas and Martin, 1991; Van Essen and Maunsell, 1983). Attention has been shown to selectively enhance the responses to low contrast stimuli (Martinez-Trujillo and Treue, 2002; Reynolds et al., 2000). Attention mediated selective enhancement of low contrast features is thought to aid invariant representations in higher object recognition areas downstream of V4 (Roe et al., 2012). However, such a bias in the attention-modulated feedback from V4 to upstream visual areas can disrupt the contrast-based feature extraction functions of these stages. How area V4 meets these distinct information processing demands of the visual processing hierarchy is not known. While attention can enhance V4 responses in a contrast-independent manner (response gain) under certain experimental conditions (Williford and Maunsell, 2006), an understanding of robust mechanisms of feedback from V4 that does not interfere with the contrast landscape of scene representations in early visual areas remains elusive.

One possibility is that distinct subpopulations in V4 mediate these functional demands. Indeed, the sensory cortical sheet, including area V4, is not a homogeneous piece of tissue along its depth; rather, it has a six-layered or laminar structure made up of multiple cell classes, of both excitatory and local inhibitory kind, with largely stereotypical anatomical connectivity between and within layers (Douglas and Martin, 2004). Layer 4 (the *input* layer) is the primary target of projections carrying visual information from early areas, such as V1, V2, and V3 (Felleman and Van Essen, 1991; Ungerleider et al., 2008). Visual information is then processed by local neural subpopulations as it is sent to layers 2/3 (the *superficial* layer) and layers 5/6 (the *deep* layer), which serve as output nodes in the laminar circuit (Hirsch and Martinez, 2006; Rockland and Pandya, 1979). The superficial layers feed information forward to downstream visual areas, such as IT (Borra et al., 2010; Distler et al., 1993), whereas the deep layers send feedback information to upstream early visual areas (Callaway, 1998; Gattass et al., 2014; Mehta et al., 2000; Ungerleider et al., 2008). This anatomical organization suggests distinct functional roles (D'Souza and Burkhalter, 2017), and differential attentional modulation of sensory representation among cell-class and layers-specific neural subpopulations. In support of this idea, a recent study of simultaneous depth recordings in visual area V4 has shown layer-specific attentional modulation of average neuronal responses, reliability of responses, and correlations between responses of pairs of neurons (Nandy et al., 2017). Therefore, to fully understand the attentional modulation of sensory computations, it is essential to investigate the modulation of sensory representation in these subpopulations. Our broad hypothesis is that the attentional modulation of contrast computations in area V4 is not homogeneous, but rather is layer- and cell-class specific and that these differences reflect the different computational demands on these subpopulations. Considering their key contribution to feedback projections to early visual areas, we specifically expect that projection neurons in the deep layers show uniform attentional modulation across all contrasts in order to minimally impact the faithful representation of contrast landscape in their target areas.

In this study, we characterized layer- and cell-class specific neural subpopulations from extracellular simultaneous laminar recordings of single neurons within area V4 of macaque monkeys performing an attention-demanding task. Using unsupervised clustering techniques on spiking properties, we distinguished five functional clusters of neurons. We distinguished layer identities – superficial, input or deep – of these neurons using features of local field potentials. To test our hypothesis, we characterized the attentional modulation of contrast response functions in these sub-populations. We interpreted our findings within a computational framework of attentional modulation of contrast responses (Reynolds and Heeger, 2009), which yielded predictions for distinct mechanistic roles of these neural subpopulations in attentive perception.

RESULTS

In the primate visual system, cortical sensitivity to features such as luminance contrast varies with the locus of spatial attention; contrast response functions (CRF) of cortical neurons are measured to quantify this dependence (Kastner and Ungerleider, 2000; Reynolds and Chelazzi, 2004; Reynolds et al., 2000). However, the laminar- and cell-class specific dependence of CRF on the attentive state is not known. Using linear array electrodes, we recorded neuronal activity from well-isolated single units, multi-unit clusters, and local field potentials (LFPs) in visual area V4 of two rhesus macaques (right hemisphere in monkey A, left hemisphere in monkey C) during an attention demanding orientation change detection task (Figure 1A, B; see Methods). We used current source density (CSD) analysis to identify different laminar compartments (*superficial*, *input*, and *deep*), and assigned isolated single units to one of the three layers (see Methods). In the main experiment, we presented a sequence of paired Gabor stimuli with different contrasts (Figure 1B); one stimulus was presented inside the receptive fields (RFs) of the recorded neurons and the other at an equally eccentric location across the vertical meridian. Attention was cued either to the stimuli within the neurons' RFs ("attend-in") or to the stimuli in the contralateral visual hemifield ("attend-away").

Attentional Modulation of Contrast Response Function

To examine the effects of attention on individual neurons, we used the method of ordinary least squares to fit each neuron's contrast responses from both attentional states to a hyperbolic ratio function (Figure 1C). This function is described by four parameters: r_{max}, c_{50}, m, and n, where r_{max} is the attainable maximum response, c_{50} is the contrast at which neuronal response is half-maximal, m is the baseline activity, and n describes the nonlinearity of the function. Attention effects differ considerably for individual neurons. Attention either enhances or suppresses neuronal responses at different contrast levels (Figure 1D). We quantified the effect of attention on every recorded neuron by computing the attentional modulation index (AMI) using contrast responses from both attention conditions (see Methods). We saw a significant variance of AMI values at each contrast level (Figure 1E). We also examined how attention impacts the values of best-fitting parameters (Figure 1F). The mean AMIs for r_{max} and m are significantly higher than zero (Mann-Whitney U test, $p < 0.01$ for both distributions), which is consistent with previous observations in V4 (Williford and Maunsell, 2006). The same percentage change in r_{max} and m (15% increase) supports an effect of contrast independent scaling by attention. The average modulations of c_{50} and n are significantly smaller than zero (Mann-Whitney U test, $p < 0.01$ for c_{50} and $p \ll 0.01$ for n), suggesting an increased sensitivity to low contrast stimuli and a reduction in the sensitivity to contrast change, respectively. The bootstrap sampling distributions of the mean difference from 0 support the average attention effects on r_{max}, n and m (Figure 1G). These results indicate that the overall effect of attention on V4 neuron responses cannot be simply explained as selective boosting of low contrast. It is a combination of modulations in multiple parameters of the contrast response function (Figure 1F, G).

Classification of Single Units Using Electrophysiological Features

To investigate whether attention modulates different classes of neurons uniformly or differentially, we characterized classes of single units based on two electrophysiological properties extracted from extracellular recordings: the peak-to-trough duration (PTD) and the local variation (Lv). Properties of the action potential waveform, especially the PTD, have been extensively used to classify neurons into narrow- (putative inhibitory) and broad-spiking (putative excitatory) cells (Constantinidis and Goldman-Rakic, 2002; Diester and Nieder, 2008; Hussar and Pasternak, 2009; Johnston et al., 2009; Kaufman et al., 2010; Mitchell et al., 2007; Wilson et al., 1994). The shapes of average spiking waveform for all single units in our data were also highly variable (Figure 2A). We exploited the information structure in the entire waveforms by applying principal component analysis (PCA). The correlation pattern between the first two components of the PCA (cumulative percentages of explained variance: 59.62%, 83.10%) supported the idea that neurons can be separated into meaningful clusters by waveform shape measures (Figure 2B). The clusters generated by neurons' PTDs in the PCA component space were minimally overlapped (Supp. Figure 2E). Therefore, we chose PTD instead of PCA components as one of the classification features for further analysis since the PTD is more interpretable.

Firing variability measures have been previously used as an additional electrophysiology-based dimension along which neurons have been found to be separable (Anderson et al., 2011; Ardid et al., 2015; Degenetais et al., 2002). We used Lv, a measure that effectively characterizes neurons' intrinsic spiking, and controls the effect of transient variations in firing rates (Shinomoto et al., 2003) (see Methods). To achieve stable classification of single units across attention conditions, we verified that Lv was not significantly modulated by attentional states (Figure 2C).

We used a meta-clustering analysis based on the *k*-means clustering algorithm (see Methods) in the two-dimensional space of PTD and Lv, and identified five clusters of isolated single units (Figure 2D) (Ardid et al., 2015; Hartigan and Wong, 1979). The five-cluster result was picked because it was the largest set of distinct cell classes that characterized a majority (99.7%) of single units in the dataset (Supp. Figure 2A). Narrow-spiking cells become a cluster by themselves, while those classified as broad-spiking cells (Mitchell et al., 2007; Nandy et al., 2017) are split into four clusters. Based on the average PTD and Lv of each cluster, we termed these five clusters as Narrow, Medium Regular, Medium Bursty, Broad Regular, and Broad Bursty.

We validated our classification results using several methods (see Methods). First, we gathered additional support for the meta-clustering based number of clusters by applying a data-driven approach based on a novel form of cross-validation (Fu and Perry, 2020). The method incorporates clustering results from the unsupervised algorithm into its supervised training of linear classifiers to produce cross-validation errors (see Methods). The five-cluster result showed the lowest cross-validation error (Supp. Figure 2B). Second, we validated the stability of the clustering result by bootstrap subsampling analysis (Hennig, 2007). The Jaccard similarity, averaged across subsamples, is a measure of each cluster's robustness regarding its sensitivity to the amount of data. All clusters in the five-cluster result had average Jaccard similarities greater than 0.5, implying that clusters remained stable under subsampling (Supp. Figure 2C). A cell-wise co-clustering matrix showing the probability that each pair of neurons belongs to a same cluster across all subsamples also supported the number of clusters we chose (Supp. Figure 2D). Third, we visualized our dataset by applying nonlinear transformations: t-SNE (Hinton and Roweis, 2003) and UMAP (McInnes et al., 2018). Although these techniques are generally suited for embedding high-dimensional data for visualization in a low-dimensional space, their algorithms that enlarge the distance differences in the original dataset also make them useful for recovering well-separated clusters. When we explored the hyperparameters of both algorithms, we found that most of the five clusters were still separable in both t-SNE and UMAP space (Figure 2E; Supp. Figure 2G, H). Notably, all four non-Narrow clusters were separable, including the Medium Regular and the Medium Bursty which occupied distinct locations in the t-SNE and UMAP space (Supp. Figure 2G, H).

One of the assumptions we made to use the PTD as a clustering feature was that it captures a significant amount of the variations of neurons' spiking waveforms. We tested this assumption by clustering neurons in the principal component space of the AP waveform and comparing them with neuronal groups defined by their PTD. We divided neurons into narrow- (0-250 μs), medium- (250-350 μs), and broad-spiking (350-550 μs) groups, and found that the 3 clusters generated from the *k*-means clustering were consistent with the 3 neuronal groups defined by the spike width (Supp. Figure 2F).

The clusters differ in terms of their firing rates (Supp. Figure 2I). Notably, Narrow class neurons exhibited higher firing rates than the Broad Regular cluster when averaged across layers (mean 10.2 Hz compared to 5.6 Hz, Mann-Whitney U test, $p < 0.05$). It is in agreement with previous findings that narrow-spiking neurons, considered putative inhibitory interneurons, show higher firing rates than broad-spiking neurons, thought to be putative excitatory pyramidal cells (Connors and Gutnick, 1990; McCormick et al., 1985; Mitchell et al., 2007; Nowak et al., 2003; Povysheva et al., 2006).

Cell-Class and Layer-Specific Attentional Modulation

We next examined how attention modulates contrast responses for each cell class. We first computed the AMIs of best-fitting CRF parameters for every cell class. The pattern of modulations of CRF parameters was distinct for individual cell classes (Figure 2F). Narrow and Medium Regular cell classes showed significant positive modulations of r_{max} only, implying a contrast-independent effect by attention. On the other hand, both Broad Regular and Broad Bursty classes showed significant negative modulations of c_{50} (Figure 2F), suggesting a selective enhancement of responses to low contrast stimuli. This effect was novel to these classes and not revealed in the analysis of unclassified neurons (Figure 1G). None of the remaining cell classes – Narrow, Medium Bursty and

Medium Regular – showed a significant modulation of c_{50} by attention, an effect that matched the analysis of unclassified neurons (Figure 1G). Medium Bursty neurons showed a modulation pattern that was distinct from the ones for any of the other four cell classes: significant positive modulations of r_{max} and baseline activity, implying a pure response gain effect by attention.

To further investigate the cell-class specific attentional modulation at each contrast level, we computed the AMI as a function of contrast using CRFs from both attentional states for every single unit and then averaged AMIs across single units within a cluster (Figure 3A, left panel). We found that the AMIs of Narrow and Medium Regular classes were relatively less dependent on contrast, whereas the remaining clusters appeared to be modulated by attention in a contrast-dependent manner (Figure 3A, left panel). When averaged across all contrasts, attention positively modulated firing rates for all cell classes except the Medium Regular class (Mann-Whitney U test, $p < 0.01$ except for MR). Further, attentional modulation differed in significant ways among the non-Narrow clusters (Figure 3A, right panel). To quantify the contrast dependence of attentional modulation for each single unit, we first averaged the AMIs within the low-contrast and the high-contrast ranges with the contrast boundary set at each unit's best-fitting c_{50} parameter. We then defined the contrast dependence index (CDI) of a single unit as the difference between the two average AMIs normalized by the AMI averaged across all contrasts (see Methods). Contrast independent modulation would then result in CDI = 0, reflecting a pure scaling effect of attention on the CRF. A positive CDI would indicate a more robust attentional modulation at the low-contrast range. A negative CDI would suggest a stronger attention effect on neural responses at the high-contrast range (Figure 3B). We examined the CDI distribution within each cell class and found that the Narrow and Medium Regular classes showed small mean CDIs, and their distributions were not significantly different from zero. However, the other 3 clusters (Medium Bursty, Broad Regular, Broad Bursty) exhibited more positive CDIs (Figure 3C). These results are consistent with our findings of AMIs of CRF parameters for each cell class (Figure 2F), confirming that attention modulated Narrow and Medium Regular cell classes' responses regardless of the stimulus contrast.

On the other hand, the modulations for Medium Bursty, Broad Regular, and Broad Bursty classes were dependent on contrast and were more robust in the low-contrast range.

We further inspected the laminar profile of the attention effect and its contrast dependence for every cell class (Figure 3D, E). We excluded from our analysis clusters that contained an insufficient number of units ($n < 10$) in a layer. When averaged across contrasts, (Figure 3D, right panels), Narrow class neurons showed significant attentional modulations in the input layer, but not in the superficial or deep layer (Figure 3D, right panels, Mann-Whitney U test, $p_{superficial} = 0.79$, $p_{input} < 0.01$, $p_{deep} = 0.06$). On the other hand, Broad Regular neurons were robustly modulated by attention across all cortical layers (Figure 3D, right panels, Mann-Whitney U test, $p_{superficial} \ll 0.01$, $p_{input} \ll 0.01$, $p_{deep} \ll 0.01$). The AMI difference between these two cell classes is in agreement with the differences between narrow- and “broad”-spiking cells previously reported in these cortical layer (Nandy et al., 2017); it is important to note that the AMI patterns across layers were distinct for the other three cell classes (Figure 3D). Two key laminar patterns of contrast dependence emerged from these 5 clusters. First, the attentional modulation of the Narrow cell class was independent of contrast across all cortical layers. Second, the Broad Regular cell class exhibited a strong contrast dependence and, specifically, a significant modulation in the low-contrast range in the superficial and input layers; but its dependence on contrast was not significant in the deep layer (Figure 3E). It is important to note that at least one non-Narrow class (Medium Regular) was functionally similar to Narrow neurons in superficial and input layers. Also notably, the laminar differences did not emerge when all units in a layer were analyzed as either a single class or more conventionally as narrow vs. “broad” classes.

Laminar network mechanisms of contrast dependence of AMI across layers

We next used computational modeling to gain insights into the possible neural mechanisms underlying the layer- and cell-class specific AMI dependency on stimulus contrast. Variation in CDI across experimental paradigms has been previous observed (Martinez-Trujillo and Treue, 2002; Reynolds et al., 2000; Williford and Maunsell, 2006), and explained by paradigm-specific normalization due to attention (Reynolds and Heeger, 2009). We hypothesized that normalization mechanisms can also explain the layer-specific differences in CDI in our empirical findings (Figure 3D, E). To test this, we first interpreted our results in the context of the normalization model of attention (Reynolds and Heeger, 2009) to generate predictions about layer-specific cortical connectivity that might underlie the variations in CDI. The normalization model of attention proposes a computational principle that accounts for various attention effects on neurons' contrast response functions (Reynolds and Heeger, 2009). Normalization model assumes that the relative sizes of excitatory receptive field and suppressive field of neurons, and the 'attention field' of the experimental paradigm shape the net suppressive drive to individual neurons. The suppressive drive ultimately determines the CDI of individual neurons in a population. We thus investigated the consequences of varying the relative sizes of excitatory receptive field and suppressive field of individual neurons on attentional modulations of CRFs (see Methods). This inquiry was motivated by the observation that neuronal receptive field sizes change along the cortical depth in sensory areas (Gilbert, 1977; Sur et al., 1985; Vaiceliunaite et al., 2013), and based on the assumption that 'attention field' sizes are constant for an experimental paradigm.

We simulated the normalization model with different sizes of excitatory receptive field and suppressive field of neurons, and generated neuronal responses to different stimulus contrasts in "attend in" and "attend away" conditions (Figure 4A, top panel). We computed the AMI and the CDI for each combination of size parameters (see Methods). We find that the CDI depends both on the excitatory receptive field size and on the suppressive field size. Holding the attention field size and the stimulus size fixed, a smaller suppressive field or a smaller excitatory receptive field leads to a greater CDI of the attentional modulation (Figure 4A, middle panel). On the other hand, a larger suppressive field or a larger excitatory receptive field results in a smaller CDI (Figure 4A, middle panel). These results hold for a wide range of values of the stimulus size and the attention field size. The pattern is robust when the attention field and the stimulus are both small or large (Supp. Figure 4B, i). The results are also stable for both a linear and saturating transfer function assumption between the stimulus contrast and excitatory drive in the normalization model (Supp. Figure 4B, ii). We also computed the AMI of suppressive drive of neurons for each combination of size parameters. The CDI of model neurons is roughly proportional to the AMI of suppressive drive (Figure 4A, bottom panel). Greater the AMI of suppressive drive, stronger is the CDI of model neurons, and vice versa. Since Broad Regular neurons are putative excitatory pyramidal cells, these results suggest two possible neural mechanisms that explain the laminar profile of CDIs of Broad Regular neurons: the suppressive field size increases along the depth of V4 (Figure 4A, middle panel) or the excitatory receptive field is more extensive in the deeper layer of V4 (Supp. Figure 4C).

The normalization model predicts the AMI of the suppressive drive (Figure 4A, bottom panel) to be correlated with the CDI of neuronal responses (Figure 4A, middle panel) (Reynolds and Heeger, 2009). However, the suppressive field in the model can be implemented by various biophysical mechanisms (Carandini, 2004). One possible mechanism is shunting inhibition via lateral connections from other neurons in the cortical neighborhood (Carandini and Heeger, 1994; Carandini et al., 1997; Kouh and Poggio, 2008), in which case the receptive field of local inhibitory neurons can approximate the suppressive field. Since the average AMI of the putative inhibitory (Narrow) cluster and CDI of putative excitatory (Broad) clusters in the input and deep layers in our empirical data (Figure 3D right panels, Figure 3E) is also correlated, we further explored this mechanism mediated by local inhibitory neurons. Under this assumption, the prediction about the changes in suppressive field size down the cortical depth from the normalization model transforms into one about changes in the excitatory (E) - inhibitory (I) connectivity along the cortical depth. Similarly, the prediction about the changes in excitatory receptive field sizes down the cortical depth can also transforms into one about the changes in the E-E connectivity along the cortical depth (Gilbert and Wiesel, 1985; Hirsch and Gilbert, 1991). The layer-specificity of cortical connectivity implies different temporal signatures of neural activity across layers.

We next used a spiking network model to examine the effects of excitatory and inhibitory receptive field sizes on spike-time correlation between populations of local excitatory (E) and inhibitory neurons (I). Our spiking network model focuses on connectivity mechanisms for generating variable sizes of suppressive and excitatory receptive fields in a cortical network. The amplitude of the spike-time correlation between neurons has been shown to depend on both the connection strength and the background synaptic noise (Ostojic et al., 2009). Therefore, the spike-time correlation between neurons can be a proxy for the size of the postsynaptic neuron's receptive field. We hypothesized that a smaller receptive field of the postsynaptic neuron would make the local connections more dominant against background inputs and lead to a higher spike-time correlation between the locally connected neurons. We examined how spike-time correlations change as a function of the inhibitory or excitatory receptive field size in a conductance-based model of spiking neurons (see Methods). We set up 10 local networks or "columns" of E and I units that were interconnected in a ring formation (Figure 4B, Supp. Figure 4C). Neurons within the same column were mutually coupled, while interactions between columns were confined to excitatory connections to local E and I neurons whose strengths decayed with distance between columns. All connections occurred with a probability of 0.5. We modeled the receptive field size as the standard deviation (σ_I or σ_E) of the connection strength between columns (Figure 4B, Supp. Figure 4C). We performed simulations that generated spiking activity in response to a step input (Figure 4B, bottom panel). The spike-time correlation between local E and I populations was calculated using pooled spike trains within the same column; the resulting spike-time correlation was averaged across columns. We found that the inhibitory receptive field size has a critical impact on the spike-time correlation amplitude in such a network (Figure 4C), while the excitatory receptive field size has little effect (Supp. Figure 4C). A larger inhibitory receptive field (larger values of σ_I) leads to a lower spike-time correlation between the local E and I populations in the network (Figure 4C). This result suggests that the prediction about inhibitory receptive field sizes down the cortical depth as the basis of CDI variation of Broad Regular neurons can be tested by examining the spike-time correlation between local E and I populations within each layer.

To test this prediction in our dataset, we computed the session-averaged spike-time correlation between Narrow (putative inhibitory neurons) and Broad Regular (putative excitatory neurons) single units within each layer (see Methods). We found that the spike-time correlation amplitudes were higher in the superficial layer and the input layer than that in the deep layer (Figure 4D). We compared the spike-time correlations in the deep layer with those in either superficial or input layers, averaged within 3 different 50ms time windows. The 95% confidence interval of the mean difference between layers in either comparison was greater than 0 for the center window (Supp. Figure 4D). In accordance with our findings from the E-I network models (Figure 4C), this suggests that inhibitory neurons in the deep layer exhibit relatively broader receptive fields, which supports the prediction by the normalization model of attention (Figure 4A, middle panel). Our findings thus provide a parsimonious explanation for the layer- and cell-class specific contrast dependence of attentional modulation observed in area V4 (Figure 4E).

DISCUSSION

Spatial attention plays a critical role in sensory guided behavior. It is thought to achieve this by enhancing the responses to low contrast stimuli in mid-tier visual cortical areas such as V4. While later stages of the visual processing hierarchy are thought to benefit from this manipulation, V4 also sends feedback projections to early visual areas that use veridical representation of contrast to aid object recognition. How area V4 meets these distinct information processing demands is not known. Contrary to the simplifying assumptions of prior empirical studies, we tested the hypothesis that V4 customizes its output to different stages of the visual processing hierarchy through layer- and cell-class specific attentional modulation of contrast computations. Recent advances in experimental techniques have shown layer- and cell-class specific functional specificity of computations in the cortical circuit (Adesnik and Naka, 2018; Adesnik and Scanziani, 2010; Naka and Adesnik, 2016; Olsen et al., 2012). However, these studies have been limited to species in which higher cognitive functions, such as attention, are challenging to study. Using computational approaches on laminar neural data in area V4 of the macaque, we find that the attentional modulation of neural responses to visual luminance contrast is indeed layer- and cell-class specific. We classified neurons into five functional cell classes defined by their action potential widths and the statistics of firing variability (Figure 2D); these classes show specificity in attention effects on their contrast response functions (Figure 2F) and the contrast dependence of attentional modulation (Figure 3C). Specifically, Narrow neurons show contrast-independent response modulation across layers; Broad Regular neurons, the putative projection neurons, exhibit significant contrast dependence of attentional modulation in the superficial layers, that project to higher level visual areas, but not in the deep layers, that project to earlier visual areas (Figure 3D, E). Notably, this highly significant laminar difference was not observable without cell-class identification. These results provide the first evidence for our broad hypothesis that attentional modulation of contrast computations in the visual cortex is heterogeneous across those cell classes and layers that project to distinct stages of the visual processing hierarchy. The qualitative nature of the attention modulation of contrast in our data is not only distinct but suggests optimization for the computational demands of the target stages. Selective boosting of responses to low contrast stimuli is compartmentalized to the superficial output layers that project representations such as extended contours and object surfaces to higher areas (see Roe et al., 2012 for a review). Contrast-independent scaling of neural responses is confined to the deep output layers. Neurons in these layers project back to early visual areas that are reliant on faithful representation of luminance contrast for low-level feature extraction. We speculate that the contrast-independent attentive feedback provides a spatial boost signal to early visual areas that do not receive direct inputs from attention control centers such as the frontal eye fields (Ungerleider et al., 2008). This also aligns with the predictive coding model of object recognition, wherein V4 is a higher-level area in the object recognition hierarchy that generates predictions of lower-level activity, without corrupting the sensory landscape that is needed for error correction (Rao and Ballard, 1999).

When interpreted within the framework of the normalization model of attention (Figure 4A), the layer-specific attention modulation predicts differences in the spatial pooling of local inhibitory populations across layers. Such differences further predict a layer-specific signature of correlations between the activities of local inhibitory and putative excitatory neurons when explored in a spiking E-I network model (Figure 4B, C). We find robust evidence for differences in inhibitory spatial pooling across layers through our analyses of correlations between putative inhibitory and putative excitatory neurons in the superficial, input, and deep layers of the cortex (Figure 4D, E).

Classification of cell-types

The duration of the extracellular spike waveform has been used to distinguish putative inhibitory interneurons from putative excitatory pyramidal cells in a wide range of species and across various brain regions (Ardid et al., 2015; Bruno and Simons, 2002; Constantinidis and Goldman-Rakic, 2002; Csicsvari et al., 1999; Fox and Ranck, 1981; Frank et al., 2001; Mitchell et al., 2007; Nandy et al., 2017; Rao et al., 1999; Simons, 1978; Swadlow, 2003; Wilson et al., 1994). In terms of attention effects, narrow-spiking neurons show stronger attention-dependent increases in absolute firing rates and firing reliability than broad-spiking cells (Mitchell et al., 2007). Statistics of the firing pattern and unsupervised clustering algorithms are also effective in identifying subpopulations of neurons with distinct functional properties (Ardid et al., 2015; Compte et al., 2003; Gouwens et al., 2019; Hawken et al., 2020; Shinomoto et al., 2009). It is important to note that the clusters we distinguished based on spike width and firing variability may not correspond to neuronal classes differentiated based on morphology or protein expression patterns (Migliore and Shepherd, 2005; Tasic et al., 2018; Zeng and Sanes, 2017). Two possible correspondences exist between the Narrow neurons and interneurons, and between the Broad Regular neurons and pyramidal cells (Connors and Gutnick, 1990; McCormick et al., 1985; Nowak et al., 2003). We find significant differences in both the firing rate (Supp. Figure 2I) and the attentional modulation of firing rates (Figure 3A, D) between clusters, suggesting their different functional roles in attention-mediated visual processing. Crucially, these distinct functional roles are reflected by the differences in contrast dependence of attentional modulation.

Relation to prior studies of spatial attention in V4

Prior studies evaluating attention effects on neuronal contrast responses proposed either contrast-independent scaling of responses, termed as response gain (McAdams and Maunsell, 1999a; Morrone et al., 2002; Pestilli et al., 2009; Treue and Martinez Trujillo, 1999) or boosting of responses to low contrast stimuli, termed as contrast gain (Li and Basso, 2008; Li et al., 2008; Martinez-Trujillo and Treue, 2002; Reynolds et al., 2000) or an intermediate effect between the two (Huang and Dobkins, 2005; Williford and Maunsell, 2006). Although the overall attentional modulation of best-fitting CRF parameters in our dataset is consistent with the intermediate effect (Figure 1F, G), attention effects on individual clusters are highly variable: a mixture effect of response gain and contrast gain is observed for Broad Regular and Broad Bursty units; Medium Bursty cluster shows a response gain change; Medium Regular and Narrow neurons are only modulated in their maximum responses (Figure 2F). Furthermore, some clusters, such as Broad Regular and Broad Bursty neurons, exhibit larger attention-dependent increases in response than the population mean, especially within the low-contrast range (Figure 3A). These observations suggest that attentional modulation of firing rate for certain cell classes may be more robust than that gleaned from previous studies that averaged across the whole recorded population. These cell-class specific increases in firing rate may significantly improve the signal-to-noise ratios of individual cell classes, and therefore, act as another important contributor to the improvement of psychophysical performance due to attention in addition to reductions in correlations (Cohen and Maunsell, 2009; Mitchell et al., 2009).

Our interpretation of the normalization model

The predictions from the normalization model (NM) of attention provide one possible explanation for the diverse contrast modulation patterns across layers. NM assumes both stimulus parameters and attention condition to contribute to the normalization input to local excitatory neurons. The stimuli presented in our experiments were optimized for the recording site and did not change with attention condition, and hence are not assumed to contribute differentially to the normalization mechanism. NM also assumes the sizes of attention field of the population to contribute to the normalization input to individual neurons. The attention field in NM describes the attention gain for each neuron in the population and depends on the animal's attentional strategy employed during the experiment (Herrmann et al., 2010). The neural substrate for the attention field is unspecified in the NM, but we assumed the attention field to be constant across the cortical depth since the data was collected using a fixed experimental paradigm. However, given a lack of the biophysical mechanism underlying attentional modulation, our understanding of the attention field may be subject to future revision. The extent of excitatory receptive field, also termed as the stimulation field, in the NM can be mediated by various cortical connectivity patterns. While we explored a lateral pooling mechanism as the determinant of the receptive field extent of neurons, innervation specificity of feedforward synaptic input could be an alternative mechanism (Bruno and Simons, 2002; Hubel and Wiesel, 1962).

The variation in contrast dependence of attentional modulation observed across layers and cell classes (Figure 3D, E) in our data is explained by the NM in a most parsimonious way via the variability of the suppressive field size (Figure 4). However, the NM is agnostic to the neural machinery dedicated to the formation of neuronal tunings or the implementation of attentional modulation. To explore the implications of its field size predictions on spike-time correlations in a biophysical model, we considered the model's stimulation field as the receptive field of putative excitatory projection neurons in a column, and its suppressive field as the receptive field of local inhibitory interneurons.

We implemented a spiking network model to relate the NM's predictions of variable suppressive field sizes to variations in spike-time correlations in our data. It is important to note that our model is not a spiking network implementation of the entirety of attention computations described by the NM. The suppressive field in NM, which mediates divisive normalization, is a computation that can be can be implemented through a variety of mechanisms (see Reynolds and Heeger, 2009 for review). We chose one of the candidate suppression mechanisms – pooling of lateral inputs by local inhibitory interneurons (Carandini and Heeger, 1994; Carandini et al., 1997; Troyer et al., 1998). A feedforward mechanism of variable suppressive fields would yield a very similar prediction for spike-time correlations between local E and I populations. Our choice was guided by excellent agreement between the NM model AMI predictions and modulation patterns of related clusters in the input and deep layers. It is, however, important to note that in the superficial layers, putative inhibitory neurons (Narrow cluster) lack significant attention modulation in spite of robust boosting of responses to low contrast stimuli in putative excitatory neurons (Broad clusters). This does not agree with the predictions of the normalization model. There are three possible explanations for this observation: 1. Suppressive drive to broad-spiking neurons in superficial layer is not provided by local inhibitory neurons within that layer. 2. Superficial layer broad-spiking neurons inherit their contrast dependent attention modulation from the input layer. 3. Suppressive drive to broad-spiking neurons in the superficial layer is provided by non-PV local inhibitory neurons within the layer. Since PV neurons are a majority of the local interneuron population which itself occupies roughly 20% of the total neural population in the cortex, it is highly possible that our recordings did not sample the other inhibitory neuronal types. Indeed, studies from the mouse visual cortex suggest that SOM+ neurons play a key role in mediating lateral inhibition to layer 2/3 pyramidal neurons (Adesnik et al, 2012). Further studies are needed to distinguish the contributions of local vs feedforward computations to the attention effects in superficial layers.

When testing the model's predictions in our dataset, we ascribed the stimulation field to any of the non-Narrow clusters, including the Broad Regular cluster identified in our layer-specific CDI analysis (Figure 3E). We ascribed the suppressive field to the receptive field of the Narrow cluster (putative interneurons). While the experimental data for the Broad Regular cluster robustly validates the model predictions (Figure 4D), the Broad Bursty and Medium Regular classes show a comparable trend (Supp. Figure 4D). We could not perform a robust analysis for the remaining non-Narrow cell classes in a subset of layers due to a lack of sufficient experimental data (Figure 3E).

Conclusion

Attention increases the signal detection abilities of individual neurons. Whether the attention mediated firing rate variability is unchanged (McAdams and Maunsell, 1999b) or reduced (Mitchell et al., 2007), the response gain alone results in improved signal-to-noise ratio of individual neurons, and enhances the discriminability of the attended signal (McAdams and Maunsell, 1999b; Verghese, 2001). Attention mediated increases in neural responses to low- and intermediate-contrast stimuli can extend the separation between the neuron's stimulus-evoked responses and its spontaneous activity, thereby improving the neuron's sensitivity to low-contrast stimuli. There has, however, been a long-standing debate regarding the nature of interactions between attention and visual scene contrast that mediate object recognition. Previous theoretical studies have sought to resolve this based on the nature of differences in experimental paradigms (Reynolds and Heeger, 2009). Our work has exploited advanced experimental techniques to bring novel understanding of these interactions. Superficial cortical layers in area V4 that project to higher object recognition stages exhibit enhancement of low contrast stimuli. Deep layers that project to earlier visual areas exhibit contrast independent attentional scaling of neuronal responses. By identifying the compartmentalization of attention modulation among cortical layers, our study has uncovered a new dimension: the nature of interactions between attention and contrast is aligned with the demands of the visual processing hierarchy. A previous study has suggested that encoding of scene contrast and spatial attention by distinct neural populations in area V1 could fulfill its visual processing demands in the face of contrast dependent attentional feedback (Pooresmaeili et al., 2010). Our work has revealed an elegant mechanism of meeting these needs via laminar compartmentalization of attention modulation in area V4 that contributes to this feedback. Low-frequency synchrony between the thalamus and visual cortex has been suggested to guide the higher-frequency synchronization of inter-area activity that is critical to the communication of attention signals between brain areas (Saalmann et al., 2012). A contrast-independent effect of attention in the deep layer of V4 may also drive alpha rhythms of pulvino-cortical loops irrespective of stimulus conditions and maintain the transmission of attentional priorities across the cortex. Future studies are needed to test these and related hypotheses about the different functional roles of contrast-attention interactions in different cortical layers.

FIGURE CAPTIONS

Figure 1. Attentional modulation of Contrast Response Function

(A) Orientation change detection task. While the monkey maintained fixation, two oriented Gabor stimuli were flashed on and off at two locations: one within the RF overlap region of the recorded V4 column and the other at a location of equal eccentricity across the vertical meridian. The covert attention of the monkey was cued to one of the two locations. One of the two stimuli changed its orientation at an unpredictable time. The monkey was rewarded for making a saccade to the location of orientation change (95% probability of change at the cued location; 5% probability at uncued location [foil trials]). If no change happened (catch trials), the monkey was rewarded for maintaining fixation.

(B) An example trial showing the single-unit signals in the attend-in condition. The time axis is referenced to the appearance of the fixation spot. Spikes (vertical ticks) in each channel come from the single unit with the highest spike rate in this trial. The gray boxes depict stimulus presentation epochs. In this particular trial, 8 sample stimuli with different contrasts were presented, followed by a target stimulus flash with an orientation change that the monkey responded to correctly. Two different waveforms were shown for two single units.

(C) The mathematical function we used to fit neuronal contrast response functions is shown on the top. Schematics at the bottom show the effect of positive attentional modulation of each parameter on the contrast response functions.

(D) The best-fitting contrast response functions of three example neurons in “attend in” and “attend away” conditions. Mean ± SEM. Insets show the attentional modulation indices calculated as a function of contrast.

(E) The AMI as a function of contrast for each of the 255 visually responsive single units, with the three example units in (C) highlighted.

(F) Attention effects on the best-fitting parameters of the contrast response function. Each histogram plots the AMI distribution of a particular parameter across the population, with the dashed line marking the 0 modulation and the arrow with a number depicting the median AMI value. The median AMI is significantly different from zero for all 4 parameters (Mann-Whitney U test, $p < 0.01$).

(G) The mean difference of AMI from 0 for the 4 parameters are shown in the Cumming estimation plot. Mean differences are plotted as bootstrap sampling distributions. Each mean difference is depicted as a dot. Each 95% confidence interval (CI) is indicated by the ends of the vertical error bars. The faded color represents that the 95% CI include 0.

Figure 2. Classification of Single Units Using Electrophysiological Features.

(A) Mean waveforms for all 410 single units recorded. Waveforms were smoothed using spline interpolation and their heights were normalized to help compare spike widths.

(B) Distribution of all single units in the space of the first two principal components (PCs) of the waveforms. The non-Gaussian structure implies that spike shape is a viable feature for classifying single units.

(C) Histogram of the local variation AMI for all units with available local variation (n=341). The dashed line marks the 0 AMI value. The arrow depicts the median value of the distribution. The average local variation of the population is not significantly modulated by attention (Mann-Whitney U test, $p = 0.37$).

(D) *k*-means clustering of 341 single units based on PTD and spiking variability. Cell classes are named after their spiking widths (narrow, medium, broad) and their spiking patterns (regular, bursty). Single units within each range of spike width are highlighted in the component space on the top. Unclassified units are displayed as black crosses in the feature space.

(E) t-SNE embedding of the same data in (D) in a 2-dimensional space. The number at the left bottom corner of each panel represents the perplexity parameter of the t-SNE embedding.

(F) The Cumming estimation plot shows the bootstrap sampling distributions of AMIs of CRF parameters for each cell class. Distributions with CIs including 0 are displayed in faded colors. The CRF parameters were only available for visually responsive single units.

Figure 3. Contrast Dependency of AMI is Cell-Class and Layer-Specific

(A) The left panel shows the AMI of contrast responses as a function of contrast averaged across visually responsive single units in each cluster. Mean ± SEM. The black line indicates the population mean. The right panel shows the mean AMI averaged across contrast for each cluster. Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$).

(B) To quantify the contrast dependence of attentional modulation, we averaged the AMI for a single unit separately within the low-contrast range and the high-contrast range (using the c_{50} as the low- to high-contrast threshold). We then defined the contrast dependence index (CDI) as the difference between the average AMI within the low-contrast range and that within the high-contrast range, normalized by the mean AMI across the whole contrast range. The schematic shows the interpretation of different ranges of CDI in terms of the AMI.

(C) The Cumming estimation plot shows the raw data of CDIs (left) and the bootstrap sampling distribution of the mean (right) for each cell class. The plus signs are the outliers within the axis range, and the arrows depict the outliers outside the axis limit. The number of valid units for each cell class is shown on the top of the swarm plot. Distributions for cell classes with CIs inclusive of 0 are shown in faded colors.

(D) Layer-wise AMI (mean ± SEM) of contrast responses for each cell class as a function of contrast (left) or averaged across contrast (right). Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$). Cell classes that contain fewer than 10 units (including outliers) are excluded.

(E) Layer-wise CDI for five clusters of units, all units, and non-narrow units. The Cumming estimation plot shows the bootstrap sampling distribution of the mean CDI. Distributions with CIs inclusive of 0 are illustrated in faded colors. The number of units excluding outliers is shown on the top of each plot. Distributions for cell classes with CIs inclusive of 0 are shown in faded colors. For the raw data of the layer-wise CDIs, see Supp. Figure 3B.

Figure 4. Computational Models Provide A Parsimonious Explanation for the Laminar Profile of AMI Contrast Dependence

(A) Predictions from the normalization model of attention with different suppressive field sizes or different excitatory (E) receptive field sizes. The top panel shows contrast response functions for a simulated neuron in the normalization model, when attending to a stimulus within the neuron's receptive field (black curve) and when attending toward the opposite hemifield (gray curve). The orange curve represents the AMI. The black dot shows the inflection point of "attend away" responses that was used to delimit the low- and high-contrast ranges. The middle panel shows CDIs for simulated neurons as a function of the E receptive field size and the suppressive field size while holding the stimulus size and the attention field size fixed. The white rectangles depict a potential mechanism that leads to the observed variation of CDIs across layers (change in suppressive field size). The black asterisk corresponds to the model parameters used for the simulation above. The bottom panel shows AMIs of suppressive drive as a function of the E receptive field size and the suppressive field size. For both simulations, the attention field size is 30 and the stimulus size is 5. The normalization model predicts the AMI of the suppressive drive to be correlated with the CDI of neuronal responses.

(B) Simulations of a conductance-based E-I network with columnized connections. Schematics of the E-I networks corresponding to the possible mechanism in (A) are shown on top. 800 E and 200 I units were evenly distributed in 10 columns around a ring. We interpret the normalization model's suppressive field as the receptive field of inhibitory neurons in the E-I network model. E and I units from the same column are mutually coupled. We modeled I receptive field size as the standard deviations (σ_I) of E-I connections (W_{ie}) across columns (middle panel, -5 and 5 are the same column). We changed the range of E-I connections across columns (W_{ie}, shades of green) while keeping other connections the same (gray, including W_{ee}, W_{ii}, W_{ei}). At the bottom, the raster plot shows the spiking activity for all units organized by their column IDs (blue, I; red, E) in response to a step input. The box depicts a 200 ms window used for computing cross-correlations between E and I populations.

(C) Cross-correlograms between E and I populations in the same column with different I receptive field sizes. Cross-correlations were calculated using the pooled spike trains of E units and I units from the same column across 500 repeats of identical simulation and averaged across 10 columns. A larger I receptive field reduces the cross-correlation between local E and I populations. Mean ± SEM.

(D) Cross-correlograms (mean ± SEM) between Narrow and Broad Regular cell classes in the superficial, input, and deep layer. Cross-correlations were calculated using the pooled spike trains of Narrow class (putative inhibitory) neurons and Broad Regular class (putative excitatory) neurons, and were averaged across sessions. The arrows mark 3 time intervals during which we averaged the cross-correlations and compared the mean differences between the superficial (or input) and the deep layers. Asterisk: The mean difference of cross-correlations in the center interval (-75 ms to 75 ms) has a 95% CI above 0. For the estimation plot, see Supp. Figure 4B.

(E) Proposed E-I networks in V4 accounting for the layer-wise CDI variations. The empirical data and the model simulations imply a larger inhibitory pooling size in the deep layer than those in the superficial and input layer. The arrows depict the canonical information flow pathways in a columnar circuit.

Supp. Figure 2. Validations for the Single-unit Classification and the AMI of CRF Parameters for Each Cell Class

(A) Percentage of classified neurons in the total sample as a function of the number of clusters (*k*) input to the *k*-means clustering algorithm. The 5-cluster result was able to identify the largest set of distinct clusters while classifying most of the units.

(B) Cross-validation (CV) errors for different numbers of clusters. The 5-cluster result shows the lowest CV error.

(C) The minimum Jaccard index across clusters for each *k* from the subsampling analysis. The analysis was applied to neuronal data from either of the two attention conditions or to combined data. Clusters that have Jaccard indices above 0.5 are considered as stable.

(D) The cell-wise co-clustering matrix showing the probability of single units belonging to the same cluster in the subsampling analysis.

(E) In the principal component space of spike shape, we colored single units based on their spike width range (open circles; narrow, medium, broad). The clusters generated from the peak-to-trough duration were minimally overlapped.

(F) In the principal component space of spike shape, we colored single units either based on their spike width range (open circles; narrow, medium, broad) or by running the *k*-means clustering algorithm with the first 2 PCs (closed circles). The clusters generated from the *k*-means clustering match the ones grouped by the peak-to-trough duration, suggesting that peak-to-trough duration is an efficient measure to capture the variance in spike shapes.

(G and H) Embedding the data used for the *k*-means clustering in a 2-dimensional space using t-SNE (G) or UMAP (H).

(I) Mean firing rate for visually responsive single units split by cell class or by layer. Neuronal firing rates were calculated from stimulus flashes with the highest common contrast across two monkey experiments in the "attend away" condition. The number of single units within each cluster is shown. In each layer, we only analyzed clusters containing more than 10 single units. Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$). Mean ± SEM.

(J) The Cumming estimation plot shows the raw data (left) of AMIs of best-fitting CRF parameters and the bootstrap sampling distribution of each cell class's mean (right).

Supp. Figure 3. Raw Data of AMIs and CDIs of AMI for Each cell class

(A) The AMI of firing rate as a function of contrast for single units within each cell class.

(B) The raw data of cluster-wise CDIs of AMI within each layer. The plus signs are the outliers within the axis range, and the arrows depict the outliers outside the axis limit. The number of valid units for each cell class is shown on the top of the swarm plot.

(C) Layer-wise AMI (mean ± SEM) for all units, Narrow unit, and non-narrow units as a function of contrast (left) or averaged across contrast (right). Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$).

Supp. Figure 4. Normalization Model of Attention and CCG Analyses Between Cell Classes

(A) The structure of the normalization model of attention. The left panel shows a pair of orientated grating stimuli with identical contrasts, acting as input to the model. The central black dot indicates the fixation point. The dashed red circle indicates the receptive field of the model neuron centered on the grating stimulus. The stimulus drive shown in the middle panel is a collection of neural activity driven by the stimuli. Neurons are arranged based on their receptive field center (horizontal position) and orientation preference (vertical position). The values of the stimulus drive are shown by brightness. The top panel shows the attention field as a function of the receptive field center and the orientation preference. In this case, the attention is guided to the right stimulus position and does not vary with orientation. Gray areas indicate values of 1, and white areas indicate values greater than 1. The suppressive drive at the bottom is calculated from the point-by-point product of the stimulus drive and the attention field and then pooled over space and orientation according to the suppressive field size. The stimulus drive is multiplied by the attention field and then divided by the suppressive field to generate the output firing rates of model neurons (right panel).

(B) i, CDIs for simulated neurons in the normalization model with different stimulus sizes and attention field sizes. In each panel, we vary the E receptive field size relative to the attention field size (x-axis), and the suppressive field size relative to the E receptive field size (y-axis). The pattern of CDI holds for a range of values of stimulus size and attention field size. ii, CDIs for simulated neurons in the normalization model with different types of inputs. We changed the stimulus drive input to the normalization model to have either a nonlinear or an attention-modulated contrast response function. We tested both the response gain (10% increase in overall response) and the contrast gain (1% of increase in detected contrast) effects. For these simulations, the attention field size is 30 and the stimulus size is 5. The pattern of CDI holds for different types of inputs.

(C) Changes in E receptive field size (white box) can also lead to the variation of CDIs across layers (left panel). We tested this hypothesis in the E-I network by adjusting the standard deviation of between-column E-E connections (W_{ee}) from narrow (green) to broad (orange) while keeping other connections the same (gray, including W_{ee}, W_{ii}, W_{ie}) (middle panel). Cross-correlograms between E and I populations in the same column suggest that different E receptive field sizes have little impact on the spike-time correlations of local neural activity across layers (right panel).

(D) Cross-correlograms (mean ± SEM) between Narrow and 3 other cell classes in the superficial, input, and deep layer. Cross-correlations were calculated using the pooled spike trains of Narrow class and the other cell class (Board Bursty, Medium Regular, or Medium Bursty) and were averaged across sessions.

(E) The Cumming estimation plot shows the mean difference for cell-class specific comparisons of average cross-correlations between the superficial (*Super.*) and deep layers or between the input and deep layers. We picked 3 time intervals to compute the average cross-correlations (rows). The raw data of average cross-correlations is plotted on the left in each panel. Each mean difference between layers is plotted on the right as a bootstrap sampling distribution.

METHODS

Attention Task and Electrophysiological Recording:

Well-isolated single units were recorded from area V4 of two rhesus macaques during an attention-demanding orientation change detection task (Figure 1A). The task design and the experimental procedures are described in detail in previous studies (Nandy et al., 2019; Nandy et al., 2017). While the monkey maintained fixation, two oriented Gabor stimuli were flashed on for 200 ms and off for variable intervals (randomly chosen between 200 and 400 ms). The contrast of each stimulus was randomly chosen from a uniform distribution of 6 contrasts (c = [10%, 18%, 26%, 34%, 42%, and 50%]). One of the stimuli was located at the receptive field overlap region of the recorded neurons and the other at an equally eccentric location across the vertical meridian. At the beginning of a block of trials, the monkey was spatially cued to covertly attend to one of the two spatial locations using instruction trials in which only one stimulus was presented. One of the two stimuli changed in orientation at an unpredictable time (minimum 1s, maximum 5s, mean 3s). The monkey was rewarded for making a saccade to the location of orientation change. 95% of the orientation changes occur at the cued location, and 5% occur at the uncued location (foil trials). We observed impaired performance and slower reaction times for the foil trials, suggesting that the monkey was indeed using the spatial cue to perform the task. The difficulty of the task was controlled by changing the degree of orientation change (randomly chosen from the following: 1°, 2°, 3°, 4°, 6°, 8°, 10°, and 12°). If no change occurred before 5 s, the monkey was rewarded for holding fixation (catch trial, 13% of trials).

While the monkey was performing the attention task, we used artificial dura chambers to facilitate the insertion of 16-channel linear array electrodes ("laminar probes", Plexon, Plexon V-probe) or single tungsten microelectrodes (FHC Inc) into cortical sites near the center of the prelunate gyrus. Neuronal signals were recorded, filtered, and stored using the Multichannel Acquisition Processor system (Plexon). Neuronal signals were classified as either isolated single units or multiunit clusters by the Plexon Offline Sorter program. For the data collected from linear array electrodes, we used current source density analysis (Mitzdorf, 1985) to identify the superficial (Layers 1-3), input (Layer 4), and deep (Layers 5 and 6) layers of the cortex based on the second derivative of the flash-triggered LFPs (Bollimunta et al., 2008; Schroeder and Lakatos, 2009; Schroeder et al., 1998; Nandy et al., 2019; Nandy et al., 2017). Cell bodies of single units with bi-phasic action potential waveforms were assigned to the same layer in which the electrode channel was situated during recordings. Units that had tri-phasic waveforms or other shapes were excluded from analyses. Extracellular data were collected over 32 sessions (23 sessions in monkey A, 9 in monkey C) using linear array electrodes and 42 sessions (24 sessions in monkey A, 18 in monkey C) using single tungsten electrodes, yielding 410 single units in total (337 units using linear array electrodes and 73 units using single tungsten electrodes). Unit yield per session was considerably higher in monkey C than monkey A, resulting in a roughly equal contribution of both monkeys toward the population data.

Contrast Response Function (CRF):

Neuronal responses were analyzed only for correctly performed trials, excluding instruction trials. We restricted all data analysis to non-target stimuli because neuronal responses to target stimuli were generally affected by the behavioral response or the reward delivery, which occurs on correct trials after the target's appearance. Moreover, the larger number of non-target stimuli compared to target stimuli provided a more reliable response strength measure. For both attention conditions, the firing rate of a single unit in response to a particular contrast was measured by counting the number of spikes within a period of 60-260 ms after stimulus onset. Its baseline firing rate in each attention condition was extracted from a 200 ms window before a stimulus flash. The mean firing rates and the standard deviations (SDs) were generated across all stimulus flashes. We considered a neuron as visually responsive if any contrast responses exceeded its baseline firing rate by 4 SDs for both attention conditions. We found that 255 of 410 single units were significantly driven by the task stimuli and had valid Lv measures (See *Analysis of Spiking Activity*).

We drew 1000 random samples of contrast responses from a normal distribution with the same mean and standard deviation as the experimental data for each visually responsive single unit. For each attention condition, we computed the CRF for each random sample by applying an ordinary least square fit to a hyperbolic ratio function:

$$r = r_{max} \cdot \frac{c^n}{c^n + c_{50}^n} + m \tag{1}$$

where r is the neuronal response, r_{max} is the maximum attainable response, c is the contrast, c_{50} is the contrast at which response is half-maximal, m is the baseline activity, and n describes the steepness of the response function and represents the neuron's sensitivity to contrast. This function has been shown to provide a good fit to contrast response functions from visual cortices in cat and macaque monkey (Albrecht and Hamilton, 1982; Williford and Maunsell, 2006). We then averaged the best-fitting CRFs across random samples to generate the mean CRF for each visually responsive single unit (Figure 1D).

Analysis of Spiking Activity:

For every single unit, the spiking variability was measured by the local variation (Lv), which quantifies the average differences between consecutive inter-spike intervals (ISIs).

$$Lv = \frac{3}{N-2}\sum_{i=1}^{N-1}\frac{(\Delta t_i - \Delta t_{i+1})^2}{(\Delta t_i + \Delta t_{i+1})^2} \tag{2}$$

where Δt_i is a given ISI and N represents the total number of spikes within the time window. The advantage of Lv over other spiking measures such as the Fano factor and coefficient of variation is that it is more robust to changes in firing rate (Shinomoto et al., 2003). We computed each unit's Lv using its spike train during a stimulus flash and averaged across all flashes (restricted to non-target stimuli).

For completely Poisson processes (where neuronal firing rates are fixed and spike times are random) the Lv is 1, whereas more regular activity takes values significantly lower than 1, and bursty spiking takes values significantly larger than 1.

Of 410 single units, we included 341 neurons with enough spikes to compute Lv for further clustering analysis.

Clustering Analysis:

We used the *k*-means clustering algorithm (Hartigan and Wong, 1979) to characterize cell classes upon the space of peak-to-trough duration (PTD) and Lv. To estimate a range of the number of clusters, we used a set of indices that evaluate the quality of clustering (Halkidi et al., 2001; Jain and Dubes, 1988; Milligan and Cooper, 1985; Vendramin et al., 2010): Rand, Mirkin, Hubert, Silhouette, Davies–Bouldin, Calinski–Harabasz, Hartigan, Homogeneity and Separation indices. We ran 50 replicates of the *k*-means clustering for different numbers of clusters, from *k* = 1 to *k* = 40. For each *k*, we selected the best replicate according to the minimum squared Euclidean distance from all cluster elements to their respective centroids. We also ran 10 identical realizations, each with a random set of initial centroids to exclude the initialization issues. We evaluated validation indices for each realization, and due to random initializations, most validation indices showed increased variability after saturation, suggesting excessive partitions in the clustering process. Based on this method, a range of 2 to 10 clusters was shown to be proper for our dataset.

We then used a meta-clustering analysis (Ardid et al., 2015) to select the most appropriate number of clusters: we ran 500 realizations of the *k*-means for each *k* and selected the best replicate from 50 replicates for each realization. After 500 realizations of each *k*, we computed the probability that pairs of neurons belonged to a same cluster. Valid clusters were identified by setting a probability threshold ($p \geq 0.9$). We considered clusters with at least five single units as reliable. We identified the most appropriate number of clusters ($k = 5$) as the largest number of reliable cell classes that classified the most neurons in the dataset (Supp. Figure 2A).

Clustering Validation:

We validated our clustering analysis in three ways. First, we applied a data-driven approach based on a form of cross-validation (Fu and Perry, 2020). We organized our data into a matrix with each row representing a single unit and each column representing a feature for clustering. We then randomly partition the rows and columns into K and L subsets, respectively. Each fold is represented by a pair (r, s) of integers, with $r \in \{1, \dots, K\}$ and $s \in \{1, \dots, L\}$. Fold (r, s) treats the rth row subset as "test" observations, and the sth column subset as "responses". The remaining $(K-1)$ row subsets are "train" observations, and the $(L-1)$ column subsets are "predictors". For our dataset, we take $K = 5$ and $L = 2$. We applied the same clustering procedures described above to the "responses" data of "train" observations to generate the cluster labels and cluster means for "train" observations. Then, we trained a linear discriminant analysis classifier with equal class priors to predict those cluster labels from the "predictors" data of "train" observations. The classifier was then applied to the "predictors" data of "test" observations to generate their predicted cluster labels as well as predicted cluster means.

The cross-validation error was then computed by averaging the squared differences between the "responses" of "test" observations and their predicted cluster means. Using such a method, we calculated the cross-validation error for each *k* (from *k* = 2 to *k* = 10) in the *k*-means clustering results (Supp. Figure 2B), and *k* = 5 showed the lowest cross-validation error.

Second, we validated the stability of our clustering analysis by subsampling analysis (Hennig, 2007). We generated 100 random subsamples containing 90% of the trials from "attend in" or "attend away" or both conditions. We computed the Lv for every single unit in subsamples. Random subsamples were then clustered by the *k*-means algorithm with *k* from 3 to 10. The Jaccard similarities were calculated between original clusters and clusters from the subsample, and the maximum was found for each original cluster. These Jaccard similarities were averaged across all subsample runs. Clusters with average Jaccard similarities below 0.5 were thought to be unstable. We reported the minimum Jaccard similarity across original clusters for each *k* (Supp. Figure 2C), and all clusters when *k* = 5 were stable. A cell-wise co-clustering matrix was also generated during this procedure (Supp. Figure 2D), and it also supported our estimation of cluster stability.

Third, we used dimensionality reduction techniques to deal with the concern that cell classes in our dataset may not be separable by the linear combinations of the two features we used as input to perform *k*-means clustering. We applied both the t-distributed stochastic neighbor embedding (t-SNE; Hinton and Roweis, 2003) algorithm and the uniform manifold approximation and projection (UMAP; McInnes et al., 2018) algorithm to our singe-unit data. Within a range of both algorithms' critical parameters, we find that the clusters from *k*-means clustering were still well separated (Supp. Figure 2G, H).

Attentional Modulation Index and its Contrast Dependency:

The attentional modulation index (AMI) of a neuron during the stimulus presentation with a specific contrast c was calculated using the best-fitting contrast response functions (r) from both attention conditions:

$$AMI(c) = \frac{r(c)^{IN} - r(c)^{AWAY}}{r(c)^{IN} + r(c)^{AWAY}} \quad (3)$$

The contrast dependence of the AMI was measured by the contrast dependence index (CDI):

$$CDI = \frac{\overline{AMI_{low}} - \overline{AMI_{high}}}{|\overline{AMI_{all}}|} \quad (4)$$

where $\overline{AMI_{low}}$ and $\overline{AMI_{high}}$ are the average AMIs within the low-contrast range and the high-contrast range, respectively. $\overline{AMI_{all}}$ is the average AMI across all contrasts. c_{50} from the best-fitting CRF during "attend away" condition delimited the range of low contrast ($c < c_{50}$) and the range of high contrast ($c \geq c_{50}$). CDI measures how the AMI of a neuron fluctuates with the contrast of the stimulus. A zero CDI indicates that the AMI is independent of the contrast of the stimulus. More robust attentional modulation at the low-contrast range leads to positive CDIs, and more potent attention effects at the high-contrast range result in negative CDIs (Figure 3B). AMI and CDI were only calculated for those visually responsive neurons whose laminar locations were identified (n = 255).

Normalization Model Simulations:

We used the normalization model of attention (Reynolds and Heeger, 2009) to explore the neural mechanisms behind the variety of attentional modulation across layers (Supp. Figure 4A). The normalization model posits that the resulting firing rate (R) of the population of simulated neurons can be produced from a function of the stimulus drive (E), the attention field (A), and the suppressive drive (S):

$$R(c; x, \theta) = \frac{A(x,\theta)E(x,\theta;c)}{S(x,\theta;c) + \sigma} \quad (5)$$

where x and θ represent the receptive field center and orientation preference of each neuron in the population. c is stimulus contrast and σ is a constant that controls the contrast gain of the neurons' response. The stimulus drive is derived from the stimulus and the stimulation field of the model neuron, which is its receptive field in the spatial and orientational space. The attention field describes the strength of attentional modulation as a function of receptive field center and orientation preference. The attentional modulation is 1 for unattended space and is greater than 1 for a small range of locations around the attended stimulus. We computed the suppressive drive by pooling the product of the stimulus drive and the attention field over spatial positions and orientations:

$$S(x,\theta;c) = s(x,\theta) * [A(x,\theta)E(x,\theta;c)] \tag{6}$$

where $s(x,\theta)$ is the suppressive field and $*$ represents convolution. The stimulus, stimulation field (excitatory receptive field), attention field, and suppressive field all had Gaussian profile in space and orientation.

For simulations in Figure 4A, the stimulus size was 5 and the attention field size was 30. The CDI pattern holds for a range of stimulus sizes and attention field sizes (Supp. Figure 4B). The excitatory receptive field size and the suppressive field size were varied according to their ratios relative to the attention field size. For a given pair of stimulus size and attention field size, we changed the ratio of the attention field size to the excitatory receptive field size from 0.5 to 3 and the ratio of the suppressive field size to the excitatory receptive field size from 1 to 6. The orientation tuning width of the excitatory receptive field was 60°, and the suppressive field was nonspecific. A baseline activity of 0.5 was added after the normalization. For each combination of parameters, the AMIs were calculated using the model neuron responses from two attention conditions. The CDIs of AMIs were computed from the average AMIs within the low-contrast range and the high-contrast range delimited by the CRF's inflection point from the "attend away" condition. For simulations in Supp. Figure 4C, we further modified the stimulus drive of the model to have either a nonlinear or an attention-modulated contrast response function. The nonlinear function was implemented as

$$r(c) = \frac{c}{c+\sigma}$$

where σ is 0.26, matching the average c_{50} of our data. We also applied either a multiplicative response gain (10% of increase in overall response) or a contrast gain (1% of increase in perceived contrast) to test the effects of different attention modulation of inputs on the model neurons' responses.

Computational Model:

We set up a conductance-based model of N_E excitatory (E) and N_I inhibitory (I) neurons with a connection probability of 0.5 (Figure 4B). Neurons were evenly divided into 10 columns or local E-I sub-networks around a ring with the following within-column synaptic weights:

$$\text{E to E}: w_{EE} = \frac{W_{EE}}{N_E};\ \text{I to I}: w_{II} = \frac{W_{II}}{N_I};\ \text{E to I}: w_{IE} = \frac{W_{IE}}{N_E};\ \text{I to E}: w_{EI} = \frac{W_{EI}}{N_I}$$

We only modeled E to I connections and E to E connections between different columns. The synaptic weights fell off with column distance following a Gaussian profile:

$$w^{ij} = \frac{W}{N_E} \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{\sigma}\right)^2\right) \tag{7}$$

where w^{ij} is the synaptic weight between two columns (w_{IE}^{ij} or w_{EE}^{ij}) and d_{ij} represents the distance from column j to column i. σ controls the pooling size of the postsynaptic inhibitory (σ_I) or excitatory (σ_E) neuron.

We simulated models of $N_E = 800$ excitatory and $N_I = 200$ inhibitory spiking units. The spiking units were modeled as Izhikevich neurons (Izhikevich, 2003) with the following dynamics:

$$\frac{dv}{dt} = 0.04v^2 + 5v + 140 - u + I \tag{8}$$

$$\frac{du}{dt} = a(bv - a) \tag{9}$$

$$if\ v \geq 30\, mV, then \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \tag{10}$$

v represents the membrane potential of the neuron and u is a membrane recovery variable. I is the current input to the neuron (synaptic and injected DC currents). The parameters a, b, c, and d determine intrinsic firing patterns and were chosen as follows:

Regular spiking excitatory units: $a = 0.02, b = 0.2, c = -65, d = 8$

Fast spiking inhibitory units: $a = 0.1, b = 0.2, c = -65, d = 2$

Presynaptic excitatory neurons generate fast (AMPA) and slow (NMDA) synaptic currents, while presynaptic inhibitory neurons generate fast GABA currents:

$$I_{syn} = \sum_i g_{AMPA}(t)(v(t) - V_{AMPA}) + \sum_j g_{NMDA}(t)(v(t) - V_{NMDA}) + \sum_k g_{GABA}(t)(v(t) - V_{GABA}) \tag{11}$$

where $V_{AMPA} = 0$, $V_{NMDA} = 0$, $V_{GABA} = -70$ are the respective reversal potentials (mV). The synaptic time course g(t) was modeled as a difference between exponentials:

$$g(t) = \frac{1}{\tau_d - \tau_r}\left[exp\left(-\frac{t - \tau_l}{\tau_d}\right) - exp\left(-\frac{t - \tau_l}{\tau_r}\right)\right] \tag{12}$$

where the parameters $\tau_d, \tau_r,$ and τ_l are the decay, rise, and latency time constants with the following values (Brunel and Wang, 2003): AMPA: $\tau_d = 2$ ms, $\tau_r = 0.5$ ms, $\tau_l = 1$ ms; NMDA: $\tau_d = 80$ ms, $\tau_r = 2$ ms, $\tau_l = 1$ ms; GABA: $\tau_d = 5$ ms, $\tau_r = 0.5$ ms, $\tau_l = 1$ ms; The AMPA to NMDA ratio is 0.45 (Myme et al., 2003).

We simulated the network with a DC step current ($I_{DC} = 4$) of duration 1.2 s. Synaptic noise was sampled from a normal distribution ($I_{syn-noise} \sim \mathcal{N}(\mu = 0, \sigma = 3)$). We pooled over spike trains of excitatory units and inhibitory units in each column separately and calculated the shuffled-corrected jittered cross-correlations from the two population spike trains binned at 1 ms within the 200 ms time window (800-1000 ms) after the initial transient response across 500 repeats of the simulation. Cross-correlations for different choices of σ_I or σ_E were reported as the average across columns (Figure 4C) (Harrison et al., 2007; Harrison and Geman, 2009).

Spike Train Cross-correlations:

The population cross-correlograms in Figure 4 report shuffled-corrected jittered cross-correlations (Harrison et al., 2007; Harrison and Geman, 2009). We computed the jittered cross-correlations by resampling two spike trains within a specific time window such that for each spike in the original data, a spike is chosen at random with replacement from within the same time window across trials, thus preserving the PSTH at the resolution of the jitter window. We computed the jittered cross-correlations with 4, 8, and 16 jitter windows, and the results of 8 jitter windows were shown. Shuffled cross-correlations were calculated by cross-correlating the first population spike train with the randomly permuted second population spike train. Both types of cross-correlations were averaged across trials and were further normalized by the geometric mean of the two spike trains' firing rates and a triangular function that corrects for the amount of overlap for the different lags. The normalized shuffled cross-correlation was then subtracted from the normalized jittered cross-correlation to produce the shuffled-corrected jittered cross-correlation.

AUTHOR CONTRIBUTIONS

MPJ & ASN conceptualized the project. XW analyzed the data, previously collected by ASN, and performed the computational modeling. MPJ supervised the project. XW, MPJ and ASN wrote the manuscript.

ACKNOWLEDGEMENTS

This research was supported by NIH R00 EY025026 to MPJ, NARSAD Young Investigator Grant, Ziegler Foundation Grant and Yale Orthwein Scholar Funds to ASN, and by NEI core grant for vision research P30 EY026878 to Yale University.

SUPPLEMENTARY MATERIAL

Figures S2-S4

REFERENCES

Adesnik, H., and Naka, A. (2018). Cracking the Function of Layers in the Sensory Cortex. Neuron *100*, 1028-1043.

Adesnik, H., and Scanziani, M. (2010). Lateral competition for cortical space by layer-specific horizontal circuits. Nature *464*, 1155-1160.

Albrecht, D.G., and Hamilton, D.B. (1982). Striate cortex of monkey and cat: contrast response function. J Neurophysiol *48*, 217-237.

Anderson, E.B., Mitchell, J.F., and Reynolds, J.H. (2011). Attentional modulation of firing rate varies with burstiness across putative pyramidal neurons in macaque visual area V4. J Neurosci *31*, 10983-10992.

Anderson, J.C., and Martin, K.A. (2006). Synaptic connection from cortical area V4 to V2 in macaque monkey. J Comp Neurol *495*, 709-721.

Ardid, S., Vinck, M., Kaping, D., Marquez, S., Everling, S., and Womelsdorf, T. (2015). Mapping of Functionally Characterized Cell Classes onto Canonical Circuit Operations in Primate Prefrontal Cortex. The Journal of Neuroscience *35*, 2975.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., and Malach, R. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. J Neurophysiol *87*, 3102-3116.

Bisley, J.W., and Goldberg, M.E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. Science *299*, 81-86.

Bollimunta, A., Chen, Y., Schroeder, C.E., and Ding, M. (2008). Neuronal Mechanisms of Cortical Alpha Oscillations in Awake-Behaving Macaques. The Journal of Neuroscience *28*, 9976.

Borra, E., Ichinohe, N., Sato, T., Tanifuji, M., and Rockland, K.S. (2010). Cortical connections to area TE in monkey: hybrid modular and distributed organization. Cereb Cortex *20*, 257-270.

Brunel, N., and Wang, X.J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. J Neurophysiol *90*, 415-430.

Bruno, R.M., and Simons, D.J. (2002). Feedforward mechanisms of excitatory and inhibitory cortical receptive fields. J Neurosci *22*, 10966-10975.

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

Carandini, M. (2004). Receptive Fields and Suppressive Fields in the Early Visual System. Cognitive Neurosciences Iii, Third Edition *3*, 313-326.

Carandini, M., and Heeger, D.J. (1994). Summation and division by neurons in primate visual cortex. Science *264*, 1333-1336.

Carandini, M., Heeger, D.J., and Movshon, J.A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci *17*, 8621-8644.

Carrasco, M., Ling, S., and Read, S. (2004). Attention alters appearance. Nat Neurosci *7*, 308-313.

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

Compte, A., Constantinidis, C., Tegner, J., Raghavachari, S., Chafee, M.V., Goldman-Rakic, P.S., and Wang, X.J. (2003). Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. J Neurophysiol *90*, 3441-3454.

Connors, B.W., and Gutnick, M.J. (1990). Intrinsic firing patterns of diverse neocortical neurons. Trends Neurosci *13*, 99-104.

Constantinidis, C., and Goldman-Rakic, P.S. (2002). Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. J Neurophysiol *88*, 3487-3497.

Csicsvari, J., Hirase, H., Czurko, A., Mamiya, A., and Buzsaki, G. (1999). Oscillatory coupling of hippocampal pyramidal cells and interneurons in the behaving Rat. J Neurosci *19*, 274-287.

D'Souza, R.D., and Burkhalter, A. (2017). A Laminar Organization for Selective Cortico-Cortical Communication. Front Neuroanat *11*, 71.

Degenetais, E., Thierry, A.M., Glowinski, J., and Gioanni, Y. (2002). Electrophysiological properties of pyramidal neurons in the rat prefrontal cortex: an in vivo intracellular recording study. Cereb Cortex *12*, 1-16.

Diester, I., and Nieder, A. (2008). Complementary contributions of prefrontal neuron classes in abstract numerical categorization. J Neurosci *28*, 7737-7747.

Distler, C., Boussaoud, D., Desimone, R., and Ungerleider, L.G. (1993). Cortical connections of inferior temporal area TEO in macaque monkeys. J Comp Neurol *334*, 125-150.

Douglas, R.J., and Martin, K.A. (1991). A functional microcircuit for cat visual cortex. J Physiol *440*, 735-769.

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

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Paper presented at: Cereb cortex (Citeseer).

Fioravanti, M., Carlone, O., Vitale, B., Cinti, M.E., and Clare, L. (2005). A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. Neuropsychol Rev *15*, 73-95.

Fox, S.E., and Ranck, J.B., Jr. (1981). Electrophysiological characteristics of hippocampal complex-spike cells and theta cells. Exp Brain Res *41*, 399-410.

Frank, L.M., Brown, E.N., and Wilson, M.A. (2001). A comparison of the firing properties of putative excitatory and inhibitory neurons from CA1 and the entorhinal cortex. J Neurophysiol *86*, 2029-2040.

Fu, W., and Perry, P.O. (2020). Estimating the Number of Clusters Using Cross-Validation. Journal of Computational and Graphical Statistics *29*, 162-173.

Gattass, R., Galkin, T.W., Desimone, R., and Ungerleider, L.G. (2014). Subcortical connections of area V4 in the macaque. J Comp Neurol *522*, 1941-1965.

Ghose, G.M., and Maunsell, J.H. (2008). Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4. J Neurosci *28*, 5115-5126.

Gilbert, C.D. (1977). Laminar differences in receptive field properties of cells in cat primary visual cortex. J Physiol *268*, 391-421.

Gilbert, C.D., and Wiesel, T.N. (1985). Intrinsic connectivity and receptive field properties in visual cortex. Vision Res *25*, 365-374.

Gouwens, N.W., Sorensen, S.A., Berg, J., Lee, C., Jarsky, T., Ting, J., Sunkin, S.M., Feng, D., Anastassiou, C.A., Barkan, E.*, et al.* (2019). Classification of electrophysiological and morphological neuron types in the mouse visual cortex. Nat Neurosci *22*, 1182-1195.

Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2001). On Clustering Validation Techniques. Journal of Intelligent Information Systems *17*, 107-145.

Harrison, M., Amarasingham, A., and Geman, S. (2007). Jitter methods for investigating spike train dependencies. Computational and Systems Neuroscience Abstracts III-17.

Harrison, M.T., and Geman, S. (2009). A rate and history-preserving resampling algorithm for neural spike trains. Neural Comput *21*, 1244-1258.

Hartigan, J.A., and Wong, M.A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society Series C (Applied Statistics) *28*, 100-108.

Hawken, M.J., Shapley, R.M., Disney, A.A., Garcia-Marin, V., Henrie, A., Henry, C.A., Johnson, E.N., Joshi, S., Kelly, J.G., Ringach, D.L.*, et al.* (2020). Functional Clusters of Neurons in Layer 6 of Macaque V1. J Neurosci *40*, 2445-2457.

Hennig, C. (2007). Cluster-wise assessment of cluster stability. Computational Statistics & Data Analysis *52*, 258-271.

Herrmann, K., Montaser-Kouhsari, L., Carrasco, M., and Heeger, D.J. (2010). When size matters: attention affects performance by contrast or response gain. Nat Neurosci *13*, 1554-1559.

Hinton, G.E., and Roweis, S.T. (2003). Stochastic neighbor embedding. Paper presented at: Advances in neural information processing systems.

Hirsch, J.A., and Gilbert, C.D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. J Neurosci *11*, 1800-1809.

Hirsch, J.A., and Martinez, L.M. (2006). Laminar processing in the visual cortical column. Curr Opin Neurobiol *16*, 377-384.

Huang, L., and Dobkins, K.R. (2005). Attentional effects on contrast discrimination in humans: evidence for both contrast gain and response gain. Vision Res *45*, 1201-1212.

Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol *160*, 106-154.

Hussar, C.R., and Pasternak, T. (2009). Flexibility of sensory representations in prefrontal cortex depends on cell type. Neuron *64*, 730-743.

Izhikevich, E.M. (2003). Simple model of spiking neurons. IEEE Trans Neural Netw *14*, 1569-1572.

Jain, A.K., and Dubes, R.C. (1988). Algorithms for clustering data (Prentice-Hall, Inc.).

Johnston, K., DeSouza, J.F., and Everling, S. (2009). Monkey prefrontal cortical pyramidal and putative interneurons exhibit differential patterns of activity between prosaccade and antisaccade tasks. J Neurosci *29*, 5516-5524.

Kastner, S., and Ungerleider, L.G. (2000). Mechanisms of visual attention in the human cortex. Annu Rev Neurosci *23*, 315-341.

Kaufman, M.T., Churchland, M.M., Santhanam, G., Yu, B.M., Afshar, A., Ryu, S.I., and Shenoy, K.V. (2010). Roles of monkey premotor neuron classes in movement preparation and execution. J Neurophysiol *104*, 799-810.

Kouh, M., and Poggio, T. (2008). A canonical neural circuit for cortical nonlinear operations. Neural Comput *20*, 1427-1451.

Li, X., and Basso, M.A. (2008). Preparing to move increases the sensitivity of superior colliculus neurons. J Neurosci *28*, 4561-4577.

Li, X., Lu, Z.L., Tjan, B.S., Dosher, B.A., and Chu, W. (2008). Blood oxygenation level-dependent contrast response functions identify mechanisms of covert attention in early visual areas. Proc Natl Acad Sci U S A *105*, 6202-6207.

Martinez-Trujillo, J., and Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. Neuron *35*, 365-370.

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

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

McCormick, D.A., Connors, B.W., Lighthall, J.W., and Prince, D.A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. J Neurophysiol *54*, 782-806.

McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:180203426.

McIntyre, R.S., Kennedy, S.H., Soczynska, J.K., Nguyen, H.T., Bilkey, T.S., Woldeyohannes, H.O., Nathanson, J.A., Joshi, S., Cheng, J.S., Benson, K.M., *et al.* (2010). Attention-deficit/hyperactivity disorder in adults with bipolar disorder or major depressive disorder: results from the international mood disorders collaborative project. Prim Care Companion J Clin Psychiatry *12*, PCC.09m00861.

Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2000). Intermodal selective attention in monkeys. II: physiological mechanisms of modulation. Cereb Cortex *10*, 359-370.

Migliore, M., and Shepherd, G.M. (2005). Opinion: an integrated approach to classifying neuronal phenotypes. Nat Rev Neurosci *6*, 810-818.

Milligan, G.W., and Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. Psychometrika *50*, 159-179.

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.

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

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.

Morrone, M.C., Denti, V., and Spinelli, D. (2002). Color and luminance contrasts attract independent attention. Curr Biol *12*, 1134-1137.

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

Myme, C.I., Sugino, K., Turrigiano, G.G., and Nelson, S.B. (2003). The NMDA-to-AMPA ratio at synapses onto layer 2/3 pyramidal neurons is conserved across prefrontal and visual cortices. J Neurophysiol *90*, 771-779.

Naka, A., and Adesnik, H. (2016). Inhibitory Circuits in Cortical Layer 5. Front Neural Circuits *10*, 35.

Nandy, A., Nassi, J.J., Jadi, M.P., and Reynolds, J. (2019). Optogenetically induced low-frequency correlations impair perception. Elife *8*, e35123.

Nandy, A.S., Nassi, J.J., and Reynolds, J.H. (2017). Laminar Organization of Attentional Modulation in Macaque Visual Area V4. Neuron *93*, 235-246.

Neuchterlein, K.H., Dawson, M.E., Ventura, J., Miklowitz, D., and Konishi, G. (1991). Information-processing anomalies in the early course of schizophrenia and bipolar disorder. Schizophrenia Research *5*, 195-196.

Nowak, L.G., Azouz, R., Sanchez-Vives, M.V., Gray, C.M., and McCormick, D.A. (2003). Electrophysiological classes of cat primary visual cortical neurons in vivo as revealed by quantitative analyses. J Neurophysiol *89*, 1541-1566.

Olsen, S.R., Bortone, D.S., Adesnik, H., and Scanziani, M. (2012). Gain control by layer six in cortical circuits of vision. Nature *483*, 47-52.

Ostojic, S., Brunel, N., and Hakim, V. (2009). How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. J Neurosci *29*, 10234-10253.

Pestilli, F., Ling, S., and Carrasco, M. (2009). A population-coding model of attention's influence on contrast response: Estimating neural effects from psychophysical data. Vision Res *49*, 1144-1153.

Petersen, S.E., and Posner, M.I. (2012). The attention system of the human brain: 20 years after. Annu Rev Neurosci *35*, 73-89.

Pooresmaeili, A., Poort, J., Thiele, A., and Roelfsema, P.R. (2010). Separable codes for attention and luminance contrast in the primary visual cortex. J Neurosci *30*, 12701-12711.

Povysheva, N.V., Gonzalez-Burgos, G., Zaitsev, A.V., Kroner, S., Barrionuevo, G., Lewis, D.A., and Krimer, L.S. (2006). Properties of excitatory synaptic responses in fast-spiking interneurons and pyramidal cells from monkey and rat prefrontal cortex. Cereb Cortex *16*, 541-552.

Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci *2*, 79-87.

Rao, S.G., Williams, G.V., and Goldman-Rakic, P.S. (1999). Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. J Neurophysiol *81*, 1903-1916.

Reynolds, J.H., and Chelazzi, L. (2004). Attentional modulation of visual processing. Annu Rev Neurosci *27*, 611-647.

Reynolds, J.H., Chelazzi, L., and Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci *19*, 1736-1753.

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

Reynolds, J.H., Pasternak, T., and Desimone, R. (2000). Attention increases sensitivity of V4 neurons. Neuron *26*, 703-714.

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

Roe, A.W., Chelazzi, L., Connor, C.E., Conway, B.R., Fujita, I., Gallant, J.L., Lu, H., and Vanduffel, W. (2012). Toward a unified theory of visual area V4. Neuron *74*, 12-29.

Rolls, E.T., and Baylis, G.C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. Exp Brain Res *65*, 38-48.

Saalmann, Y.B., Pinsk, M.A., Wang, L., Li, X., and Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. Science *337*, 753-756.

Schroeder, C.E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends in Neurosciences *32*, 9-18.

Schroeder, C.E., Mehta, A.D., and Givre, S.J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. Cerebral Cortex *8*, 575-592.

Shinomoto, S., Kim, H., Shimokawa, T., Matsuno, N., Funahashi, S., Shima, K., Fujita, I., Tamura, H., Doi, T., Kawano, K.*, et al.* (2009). Relating neuronal firing patterns to functional differentiation of cerebral cortex. PLoS Comput Biol *5*, e1000433.

Shinomoto, S., Shima, K., and Tanji, J. (2003). Differences in Spiking Patterns Among Cortical Neurons. Neural Computation *15*, 2823-2842.

Simons, D.J. (1978). Response properties of vibrissa units in rat SI somatosensory neocortex. J Neurophysiol *41*, 798-820.

Sur, M., Garraghty, P.E., and Bruce, C.J. (1985). Somatosensory cortex in macaque monkeys: laminar differences in receptive field size in areas 3b and 1. Brain Res *342*, 391-395.

Swadlow, H.A. (2003). Fast-spike interneurons and feedforward inhibition in awake sensory neocortex. Cereb Cortex *13*, 25-32.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., *et al.* (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72-78.

Treue, S., and Martinez Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature *399*, 575-579.

Treue, S., and Maunsell, J.H. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. Nature *382*, 539-541.

Troyer, T.W., Krukowski, A.E., Priebe, N.J., and Miller, K.D. (1998). Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity. J Neurosci *18*, 5908-5927.

Ungerleider, L.G., Galkin, T.W., Desimone, R., and Gattass, R. (2008). Cortical connections of area V4 in the macaque. Cereb Cortex *18*, 477-499.

Vaiceliunaite, A., Erisken, S., Franzen, F., Katzner, S., and Busse, L. (2013). Spatial integration in mouse primary visual cortex. J Neurophysiol *110*, 964-972.

Van Essen, D.C., and Maunsell, J.H.R. (1983). Hierarchical organization and functional streams in the visual cortex. Trends in Neurosciences *6*, 370-375.

Vendramin, L., Campello, R.J.G.B., and Hruschka, E.R. (2010). Relative clustering validity criteria: A comparative overview. Statistical Analysis and Data Mining: The ASA Data Science Journal *3*, 209-235.

Verghese, P. (2001). Visual search and attention: a signal detection theory approach. Neuron *31*, 523-535.

Williford, T., and Maunsell, J.H.R. (2006). Effects of Spatial Attention on Contrast Response Functions in Macaque Area V4. J Neurophysiol *96*, 40-54.

Wilson, F.A., O'Scalaidhe, S.P., and Goldman-Rakic, P.S. (1994). Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. Proc Natl Acad Sci U S A *91*, 4009-4013.

Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. Nat Rev Neurosci *18*, 530-546.

Figure 1 - Wang

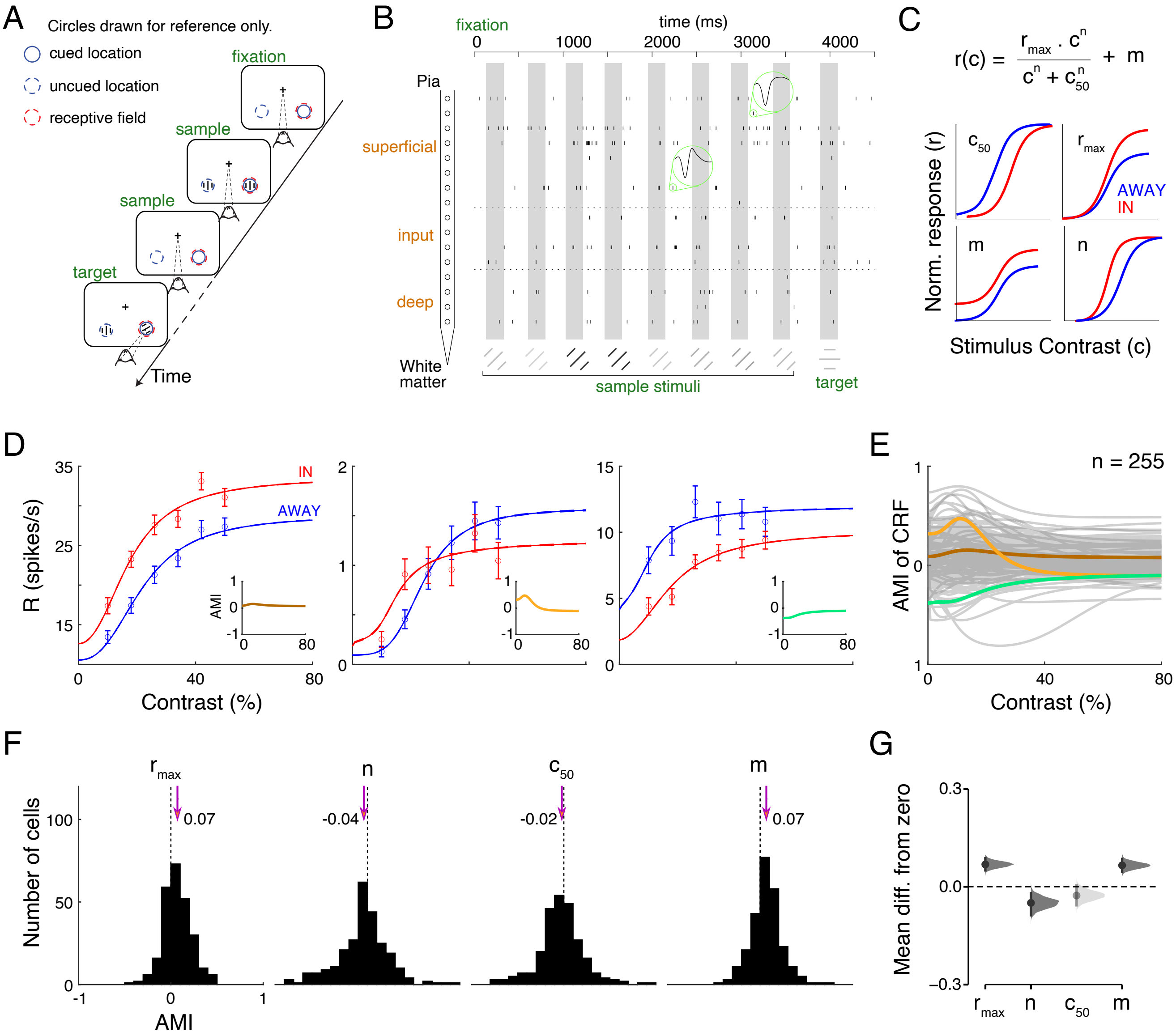


Figure 1. Attentional Modulation of Contrast Response Function

(A) Orientation change detection task. While the monkey maintained fixation, two oriented Gabor stimuli were flashed on and off at two locations: one within the RF overlap region of the recorded V4 column and the other at a location of equal eccentricity across the vertical meridian. The covert attention of the monkey was cued to one of the two locations. One of the two stimuli changed its orientation at an unpredictable time. The monkey was rewarded for making a saccade to the location of orientation change (95% probability of change at the cued location; 5% probability at uncued location [foil trials]). If no change happened (catch trials), the monkey was rewarded for maintaining fixation. Two different waveforms were shown for two single units.

(B) An example trial showing the single-unit signals in the attend-in condition. The time axis is referenced to the appearance of the fixation spot. Spikes (vertical ticks) in each channel come from the single unit with the highest spike rate in this trial. The gray boxes depict stimulus presentation epochs. In this particular trial, 8 sample stimuli with different contrasts were presented, followed by a target stimulus flash with an orientation change that the monkey responded to correctly.

(C) The mathematical function we used to fit neuronal contrast response functions is shown on the top. Schematics at the bottom show the effect of positive attentional modulation of each parameter on the contrast response function.

(D) The best-fitting contrast response functions of three example neurons in "attend in" and "attend away" conditions. Mean ± SEM. Insets show the attentional modulation indices calculated as a function of contrast.

(E) The AMI as a function of contrast for each of the 255 visually responsive single units, with the three example units in (C) highlighted.

(F) Attention effects on the best-fitting parameters of the contrast response function. Each histogram plots the AMI distribution of a particular parameter across the population, with the dashed line marking the 0 modulation and the arrow with a number depicting the median AMI value. The median AMI is significantly different from zero for all 4 parameters (Mann-Whitney U test, $p < 0.01$).

(G) The mean difference of AMI from 0 for the 4 parameters are shown in the Cumming estimation plot. Mean differences are plotted as bootstrap sampling distributions. Each mean difference is depicted as a dot. Each 95% confidence interval (CI) is indicated by the ends of the vertical error bars. The shaded color represents that the 95% CI does not include 0.

Figure 2 - Wang

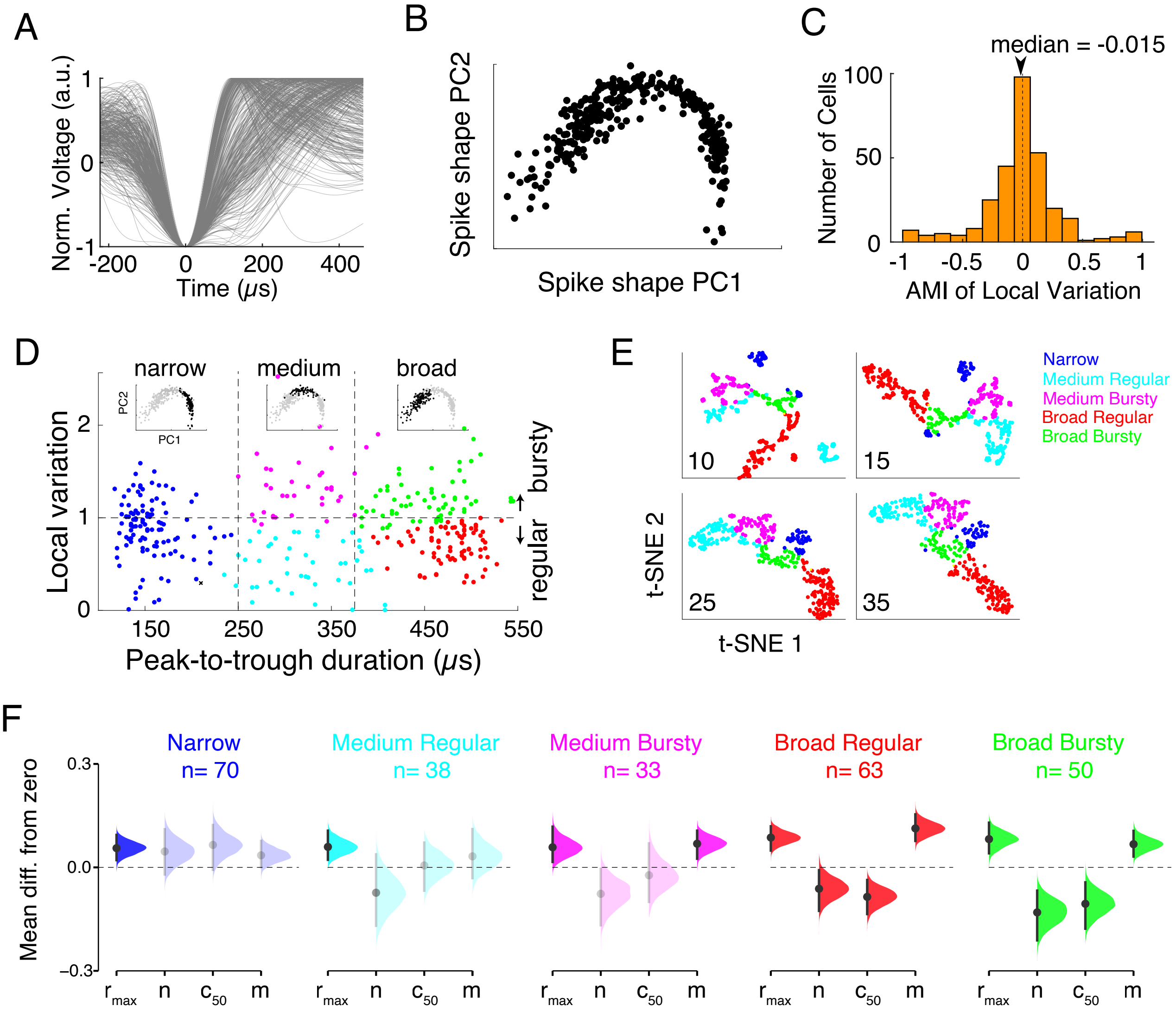


Figure 2. Classification of Single Units Using Electrophysiological Features.

(A) Mean waveforms for all 410 single units recorded. Waveform heights have been normalized to help compare spike widths.

(B) Distribution of all single units in the space of the first two principal components (PCs) of the waveforms. The non-Gaussian structure implies that spike shape is a viable feature for classifying single units.

(C) Histogram of the local variation AMI for all units with available local variation (n=341). The dashed line marks the 0 AMI value. The arrow depicts the median value of the distribution. The average local variation of the population is not significantly modulated by attention (Mann-Whitney U test, p = 0.37).

(D) *k*-means clustering of 341 single units based on PTD and spiking variability. Cell classes are named after their spiking widths (narrow, medium, broad) and their spiking patterns (regular, bursty). Single units within each range of spike width are highlighted in the component space on the top. Unclassified units are displayed as black crosses in the feature space.

(E) t-SNE embedding of the same data in (D) in a 2-dimensional space. The number at the left bottom corner of each panel represents the perplexity paramter of the t-SNE embedding.

(F) The Cumming estimation plot shows the bootstrap sampling distributions of AMIs of CRF parameters for each cell class. The CRF parameters were only available for visually responsive single units.

Figure 3 - Wang

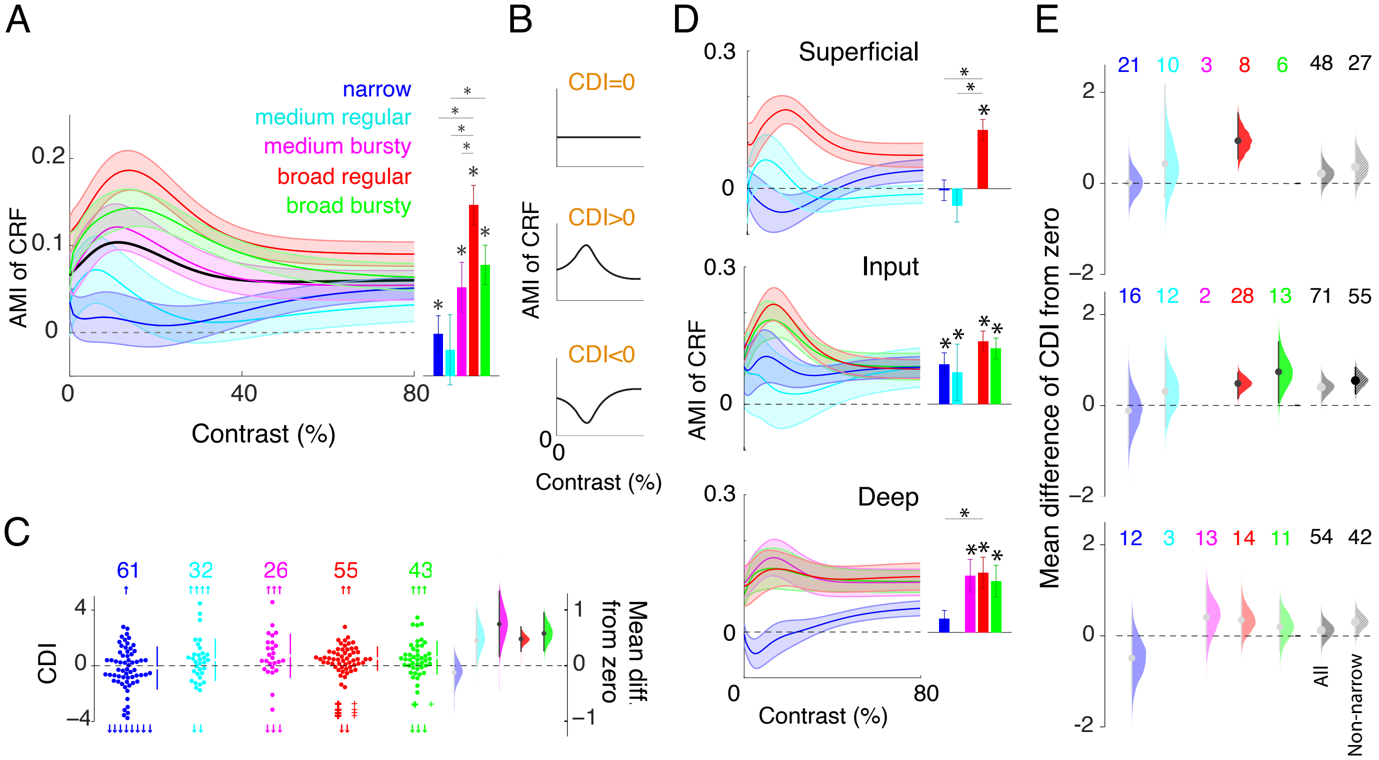


Figure 3. Contrast Dependency of AMI is Cell-Class and Layer-Specific

(A) The left panel shows the AMI of contrast responses as a function of contrast averaged across visually responsive single units in each cluster. Mean ± SEM. The black line indicates the population mean. The right panel shows the mean AMI averaged across contrast for each cluster. Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$).

(B) To quantify the contrast dependence of attentional modulation, we averaged the AMI for a single unit separately within the low-contrast range and the high-contrast range (using the c_{50} as the low- to high-contrast threshold). We then defined the contrast dependence index (CDI) as the difference between the average AMI within the low-contrast range and that within the high-contrast range, normalized by the mean AMI across the whole contrast range. The schematic shows the interpretation of different ranges of CDI in terms of the AMI.

(C) The Cumming estimation plot shows the raw data of CDIs (left) and the bootstrap sampling distribution of the mean (right) for each cell class. The plus signs are the outliers within the axis range, and the arrows depict the outliers outside the axis limit. The number of valid units for each cell class is shown on the top of the swarm plot. Distributions for cell classes with CIs inclusive of 0 are shown in faded colors.

(D) Layer-wise AMI (mean ± SEM) of contrast responses for each cell class as a function of contrast (left) or averaged across contrast (right). Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$). Cell classes that contain fewer than 10 units (including outliers) are excluded.

(E) Layer-wise CDI for five clusters, all units, and non-narrow units. The Cumming estimation plot shows the bootstrap sampling distribution of the mean CDI. Distributions with CIs inclusive of 0 are illustrated in faded colors. The number of units excluding outliers is shown on the top of each plot. For the raw data of the layer-wise CDIs, see Supp. Figure 3B.

Figure 4 - Wang

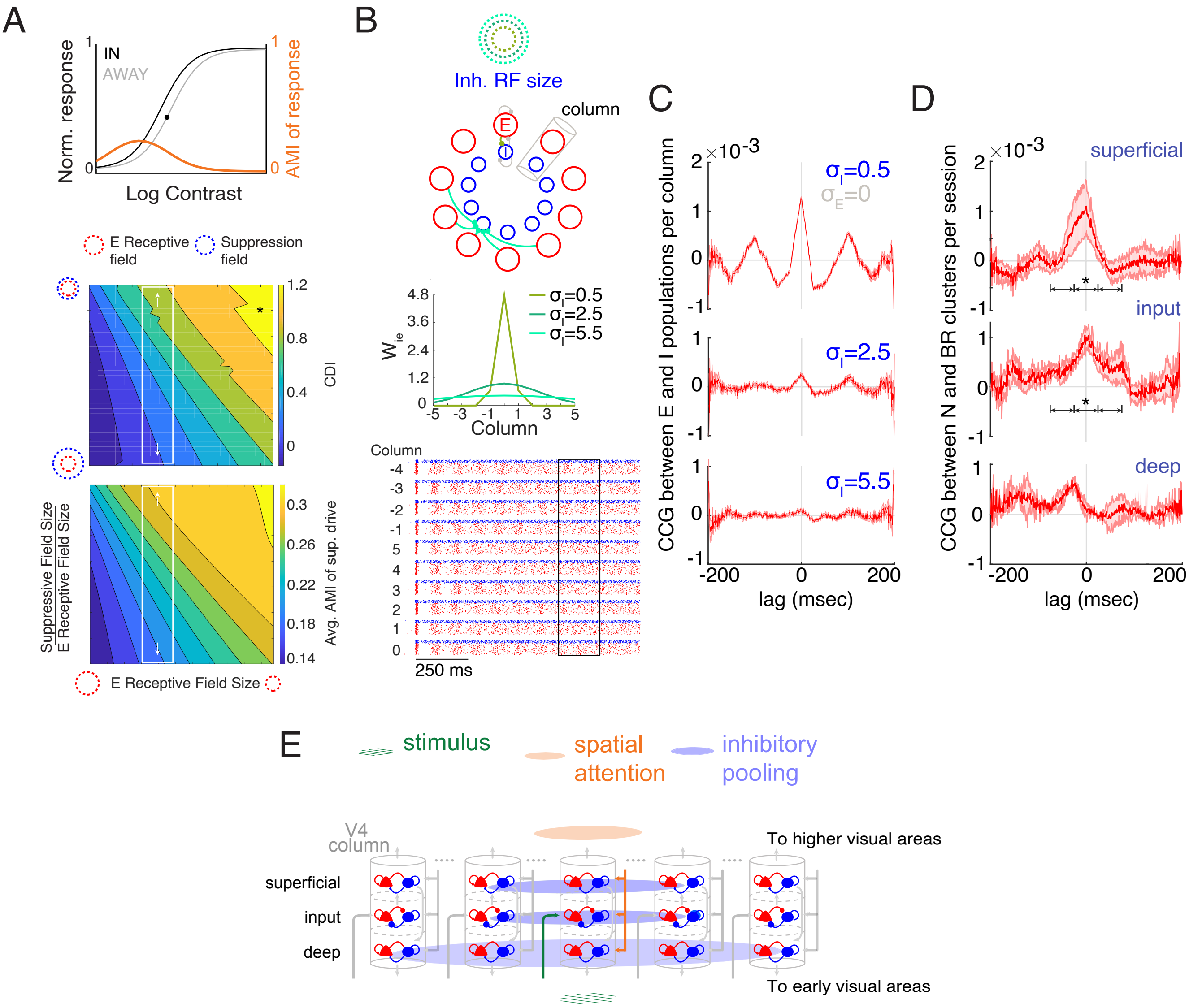
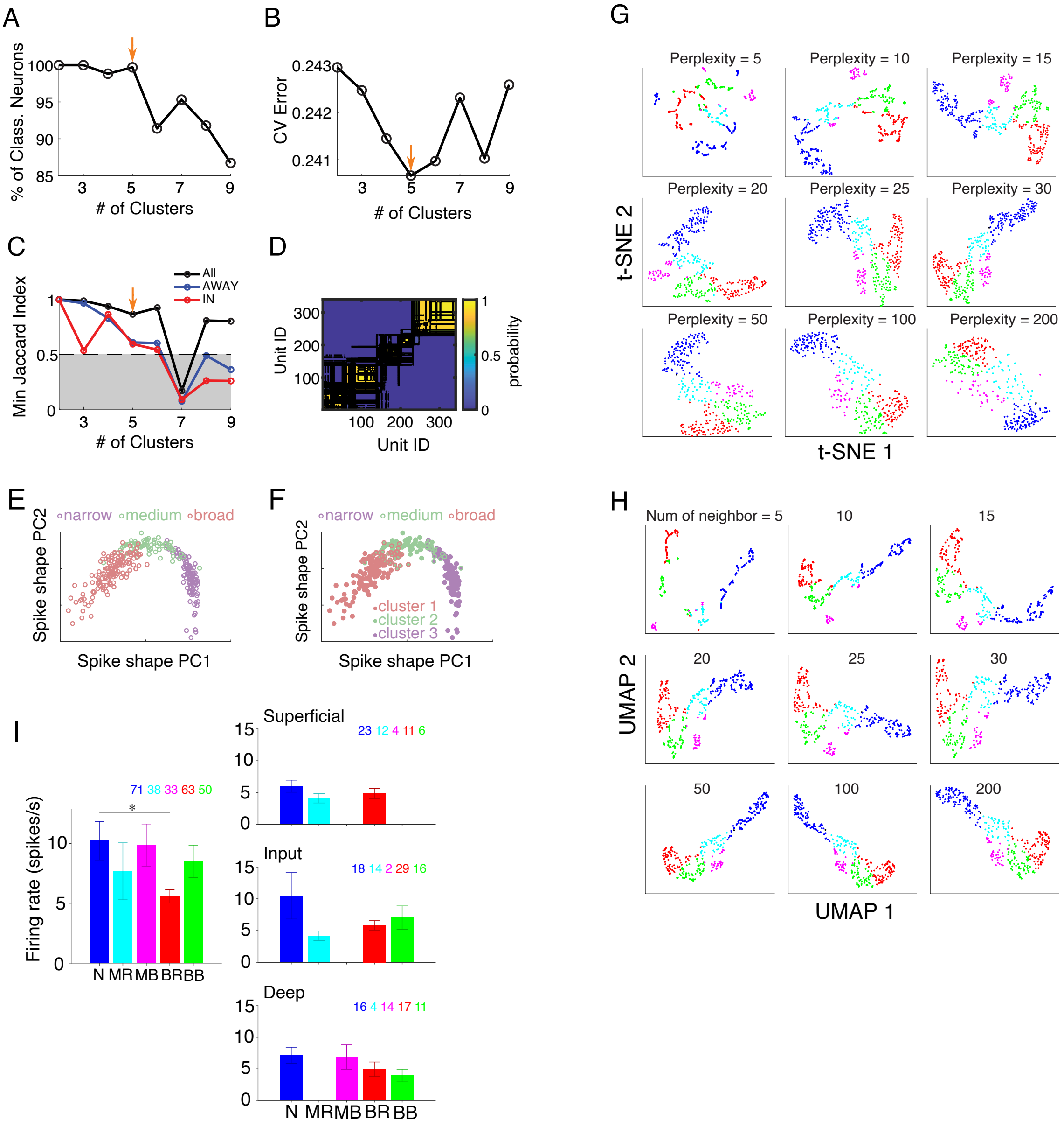


Figure 4. Computational Models Provide A Parsimonious Explanation for the Laminar Profile of AMI Contrast Dependence (A) Predictions from the normalization model of attention with different suppressive field sizes or different excitatory (E) receptive field sizes. The top panel shows contrast response functions for a simulated neuron in the normalization model, when attending to a stimulus within the neuron's receptive field (black curve) and when attending toward the opposite hemifield (gray curve). The aorange curve represents the AMI. The black dot shows the inflection point of "attend away" responses that was used to delimit the low- and high-contrast ranges. The middle panel shows CDIs for simulated neurons as a function of the E receptive field size and the suppressive field size while holding the stimulus size and the attention field size fixed. The white rectangles depict a potential mechanism that leads to the observed variation of CDIs across layers (change in suppressive field size). The black asterisk corresponds to the model parameters used for the simulation above. The bottom panel shows AMIs of suppressive drive as a function of the E receptive field size and the suppressive field size. For both simulations, the attention field size is 30 and the stimulus size is 5. The normalization model predicts the AMI of the suppressive drive to be correlated with the CDI of neuronal responses. (B) Simulations of a conductance-based E-I network with columnized connections. Schematics of the E-I networks corresponding to the possible mechanism in (A) are shown on top. 800 E and 200 I units were evenly distributed in 10 columns around a ring. We interpret the normalization model's suppressive field as the receptive field of inhibitory neurons in the E-I network model. E and I units from the same column are mutually coupled. We modeled I receptive field size as the standard deviations (σ_I) of E-I connections (W_ie) across columns (middle panel, -5 and 5 are the same column). We increased the range of E-I connections across columns (W_ie) (shades of green) while keeping other connections the same (gray, including W_ee, W_ii, W_ei). At the bottom, the raster plot shows the spiking activity for all units organized by their column IDs (blue, I; red, E) in response to a step input. The box depicts a 200 ms window used for computing cross-correlations between E and I populations. (C) Cross-correlograms between E and I populations in the same column with different I receptive field sizes. Cross-correlations were calculated using the pooled spike trains of E units and I units from the same column across 500 repeats of identical simulation and averaged across 10 columns. A larger I receptive field reduces the cross-correlation between local E and I populations. Mean ± SEM.

(D) Cross-correlograms (mean ± SEM) between Narrow and Broad Regular cell classes in the superficial, input, and deep layer. Cross-correlations were calculated using the pooled spike trains of Narrow class (putative inhibitory) neurons and Broad Regular class (putative excitatory) neurons, and were averaged across sessions. The arrows mark 3 time intervals during which we averaged the cross-correlations and compared the mean differences between the superficial (or input) and the deep layers. Asterisk: The mean difference of cross-correlations in the center interval (-75 ms to 75 ms) has a 95% CI above 0. For the estimation plot, see Supp. Figure 4B.

(E) Proposed E-I networks in V4 accounting for the layer-wise CDI variations. The empirical data and the model simulations imply a larger inhibitory pooling size in the deep layer than those in the superficial and input layer. The arrows depict the canonical information flow pathways in a columnar circuit.

Supp. Figure 2 - Wang



Supp. Figure 2. Validations for the Single-unit Classification and the AMI of CRF Parameters for Each Cell Class

(A) Percentage of classified neurons in the total sample as a function of the number of clusters (k) input to the k-means clustering algorithm. The 5-cluster result was able to identify the largest set of distinct clusters while classifying most of the units.

(B) Cross-validation (CV) errors for different numbers of clusters. The 5-cluster result shows the lowest CV error.

(C) The minimum Jaccard index across clusters for each k from the subsampling analysis. The analysis was applied to neuronal data from either of the two attention conditions or to combined data. Clusters that have Jaccard indices above 0.5 are considered as stable.

(D) The cell-wise co-clustering matrix showing the probability of single units belonging to the same cluster in the subsampling analysis.

(E) In the principal component space of spike shape, we colored single units based on their spike width range (open circles; narrow, medium, broad). The clusters generated from the peak-to-trough duration were minimally overlapped.

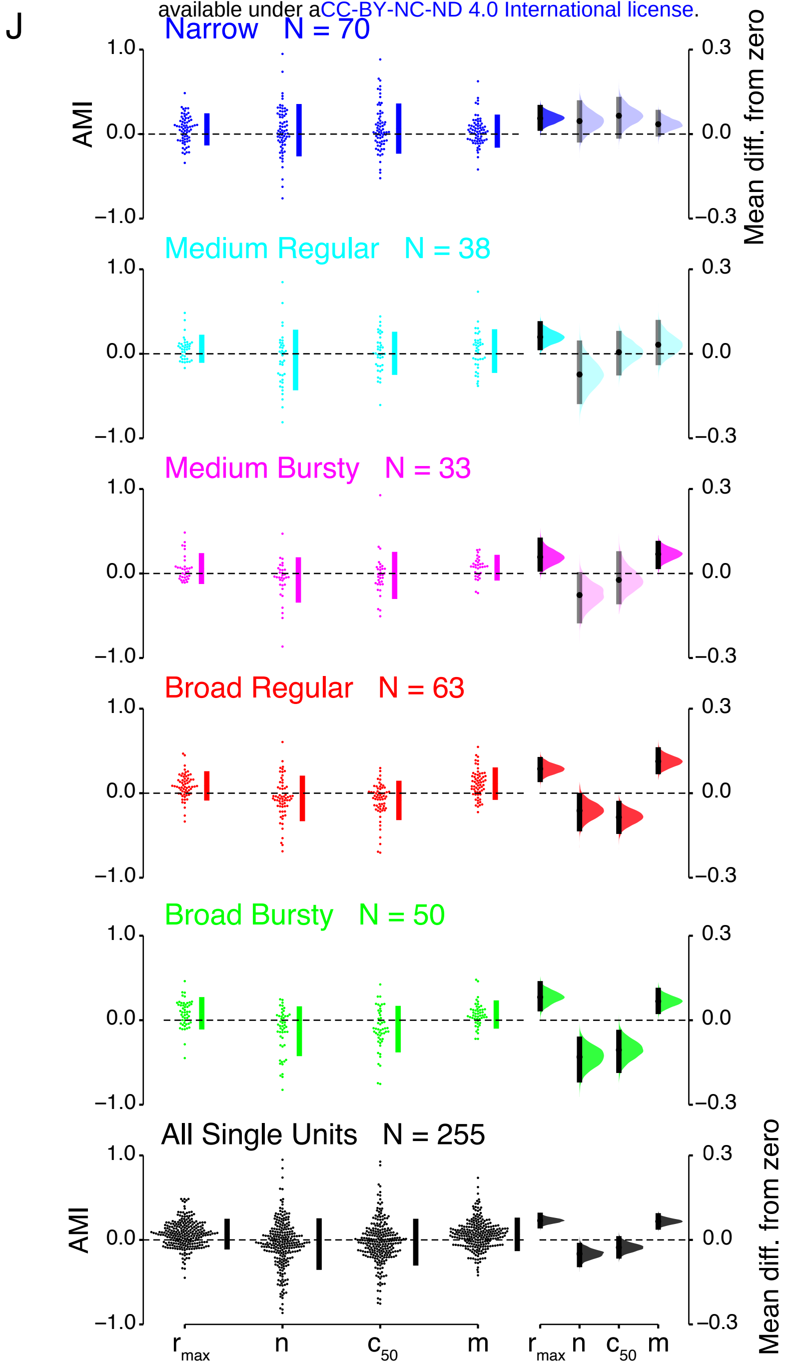
(F) In the principal component space of spike shape, we colored single units either based on their spike width range (open circles; narrow, medium, broad) or by running the k-means clustering algorithm with the first 2 PCs (closed circles). The clusters generated from the k-means clustering match the ones grouped by the peak-to-trough duration, suggesting that peak-to-trough duration is an efficient measure to capture the variance in spike shapes.

(G and H) Embedding the data used for the k-means clustering in a 2-dimensional space using t-SNE (G) or UMAP (H).

(I) Mean firing rate for visually responsive single units split by cell class or by layer. Neuronal firing rates were calculated from stimulus flashes with the highest common contrast across two monkey experiments in the "attend away" condition. The number of single units within each cluster is shown. In each layer, we only analyzed clusters containing more than 10 single units. Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$). Mean ± SEM.

(J) The Cumming estimation plot shows the raw data (left) of AMIs of best-fitting CRF parameters and the bootstrap sampling distribution of each cell class's mean (right).

Supp. Figure 2 - Wang



Supp. Figure 2. Validations for the Single-unit Classification and the AMI of CRF Parameters for Each Cell Class

(A) Percentage of classified neurons in the total sample as a function of the number of clusters (k) input to the k-means clustering algorithm. The 5-cluster result was able to identify the largest set of distinct clusters while classifying most of the units.

(B) Cross-validation (CV) errors for different numbers of clusters. The 5-cluster result shows the lowest CV error.

(C) The minimum Jaccard index across clusters for each k from the subsampling analysis. The analysis was applied to neuronal data from either of the two attention conditions or to combined data. Clusters that have Jaccard indices above 0.5 are considered as stable.

(D) The cell-wise co-clustering matrix showing the probability of single units belonging to the same cluster in the subsampling analysis.

(E) In the principal component space of spike shape, we colored single units based on their spike width range (open circles; narrow, medium, broad). The clusters generated from the peak-to-trough duration were minimally overlapped.

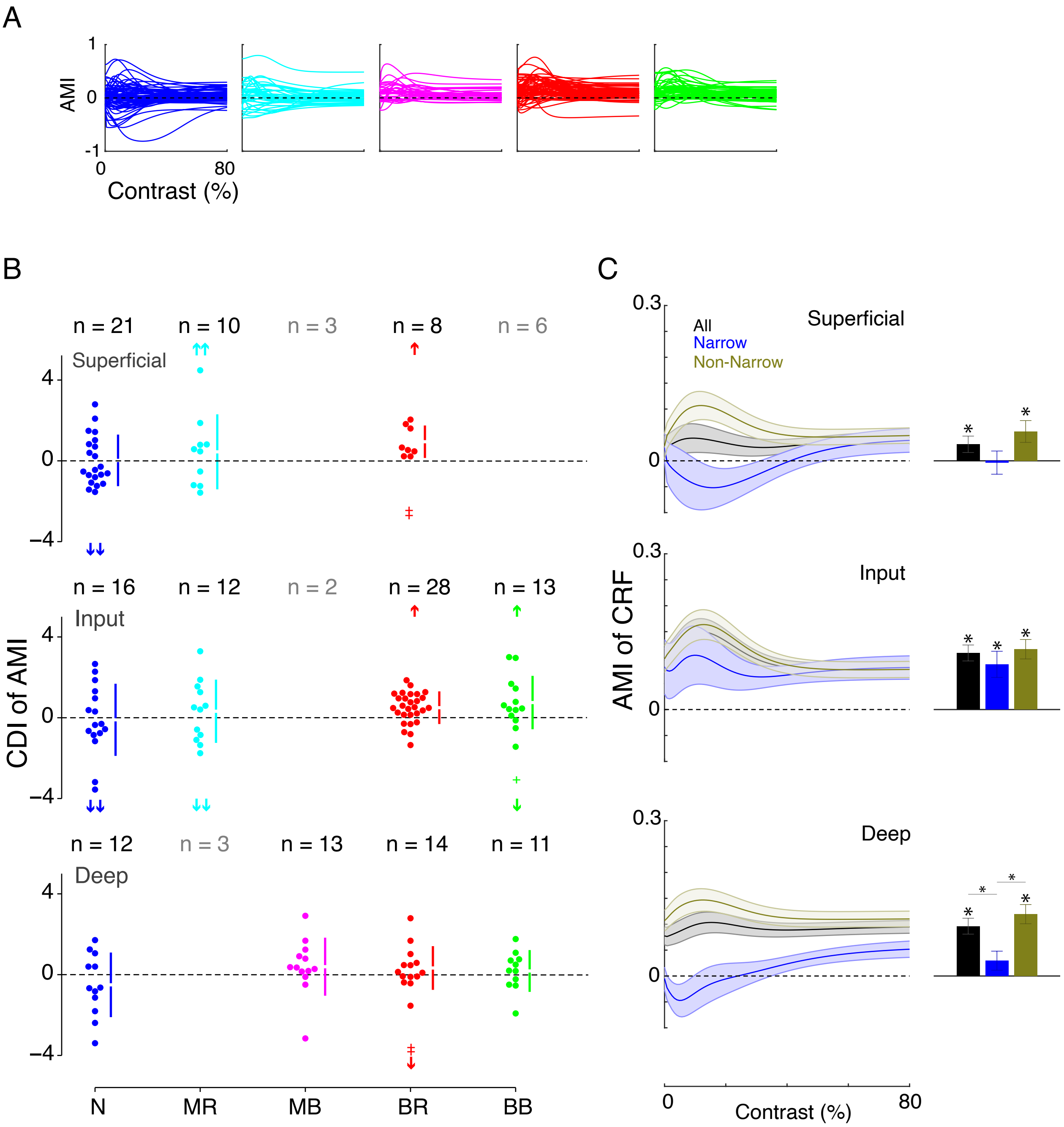
(F) In the principal component space of spike shape, we colored single units either based on their spike width range (open circles; narrow, medium, broad) or by running the k-means clustering algorithm with the first 2 PCs (closed circles). The clusters generated from the k-means clustering match the ones grouped by the peak-to-trough duration, suggesting that peak-to-trough duration is an efficient measure to capture the variance in spike shapes.

(G and H) Embedding the data used for the k-means clustering in a 2-dimensional space using t-SNE (G) or UMAP (H).

(I) Mean firing rate for visually responsive single units split by cell class or by layer. Neuronal firing rates were calculated from stimulus flashes with the highest common contrast across two monkey experiments in the "attend away" condition. The number of single units within each cluster is shown. In each layer, we only analyzed clusters containing more than 10 single units. Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$). Mean ± SEM.

(J) The Cumming estimation plot shows the raw data (left) of AMIs of best-fitting CRF parameters and the bootstrap sampling distribution of each cell class's mean (right).

Supp. Figure 3 - Wang



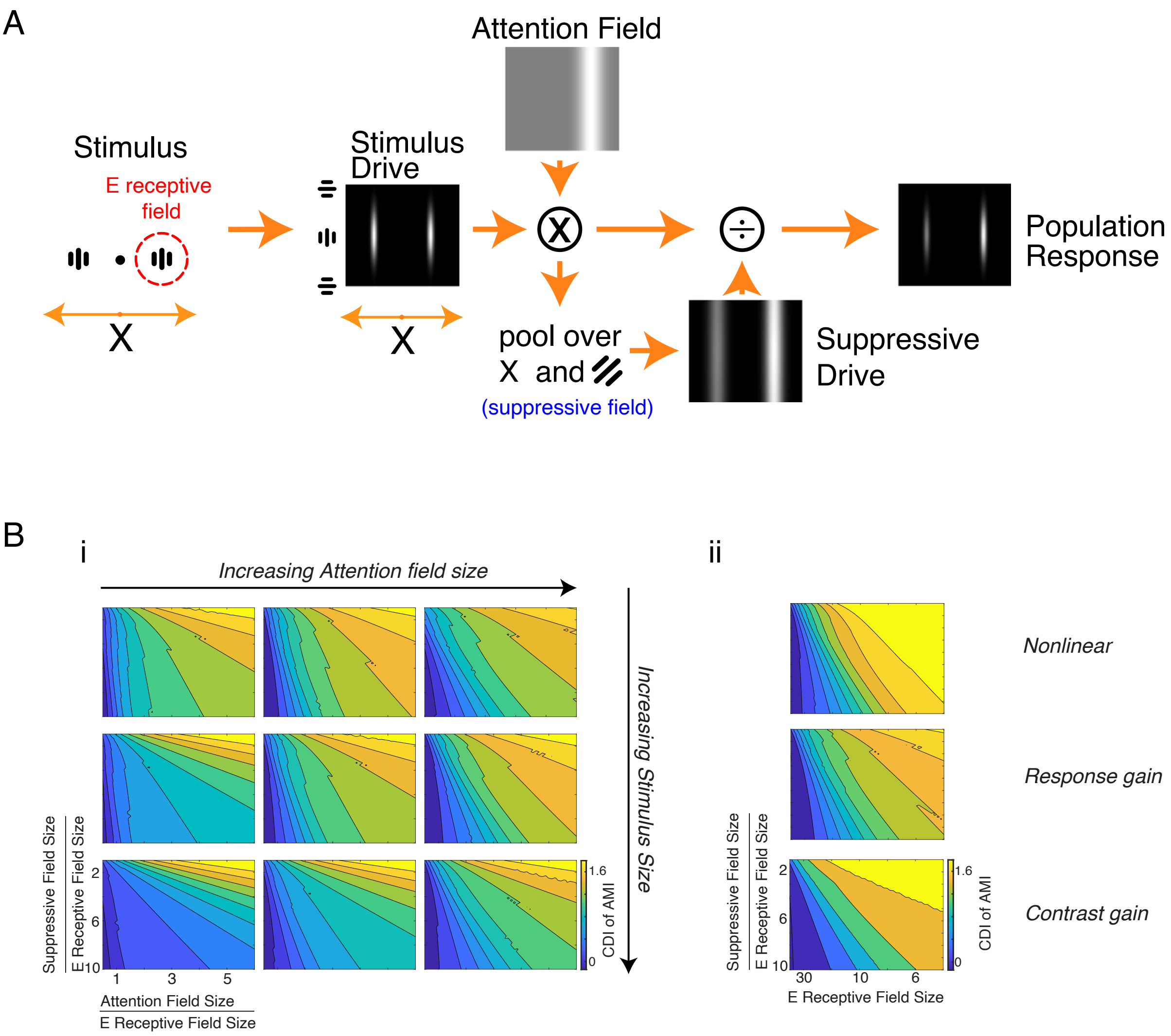
Supp. Figure 3. Raw Data of AMIs and CDIs of AMI for Each cell class

(A) The AMI of firing rate as a function of contrast for single units within each cell class.

(B) The raw data of cluster-wise CDIs of AMI within each layer. The plus signs are the outliers within the axis range, and the arrows depict the outliers outside the axis limit. The number of valid units for each cell class is shown on the top of the swarm plot.

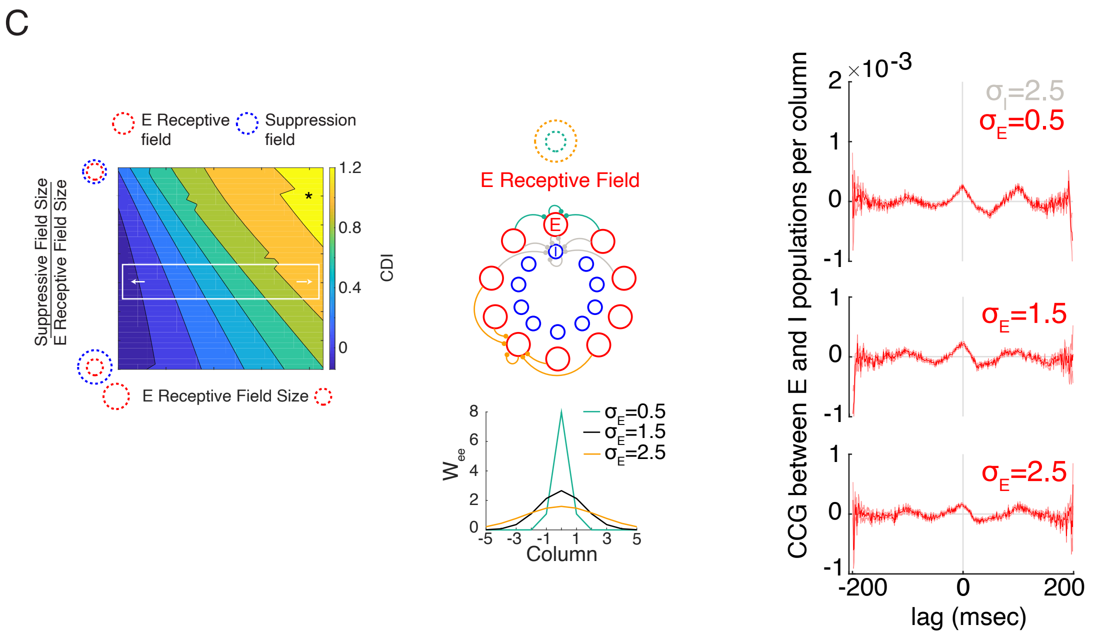
(C) Layer-wise AMI (mean ± SEM) for all units, Narrow unit, and non-narrow units as a function of contrast (left) or averaged across contrast (right). Asterisk indicates either the distribution is significantly different from zero or two distributions are significantly different (Mann-Whitney U test, $p < 0.05$).

Supp. Figure 4 - Wang



Supp. Figure 4. Normalization Model of Attention and CCG Analyses Between Cell Classes
(A) The structure of the normalization model of attention. The left panel shows a pair of orientated grating stimuli with identical contrasts, acting as input to the model. The central black dot indicates the fixation point. The dashed red circle indicates the receptive field of the model neuron centered on the grating stimulus. The stimulus drive shown in the middle panel is a collection of neural activity driven by the stimuli. Neurons are arranged based on their receptive field center (horizontal position) and orientation preference (vertical position). The values of the stimulus drive are shown by brightness. The top panel shows the attention field as a function of the receptive field center and the orientation preference. In this case, the attention is guided to the right stimulus position and does not vary with orientation. Gray areas indicate values of 1, and white areas indicate values greater than 1. The suppressive drive at the bottom is calculated from the point-by-point product of the stimulus drive and the attention field and then pooled over space and orientation according to the suppressive field size. The stimulus drive is multiplied by the attention field and then divided by the suppressive field to generate the output firing rates of model neurons (right panel).
(B) i, CDIs for simulated neurons in the normalization model with different stimulus sizes and attention field sizes. In each panel, we vary the E receptive field size relative to the attention field size (x-axis), and the suppressive field size relative to the E receptive field size (y-axis). The pattern of CDI holds for a range of values of stimulus size and attention field size. ii, CDIs for simulated neurons in the normalization model with different types of inputs. We changed the stimulus drive input to the normalization model to have either a nonlinear or an attention-modulated contrast response function. We tested both the response gain (10% increase in overall response) and the contrast gain (1% of increase in detected contrast) effects. For these simulations, the attention field size is 30 and the stimulus size is 5. The pattern of CDI holds for different types of inputs.
(C) Changes in E receptive field size (white box) can also lead to the variation of CDIs across layers (left panel). We tested this hypothesis in the E-I network by adjusting the standard deviation of between-column E-E connections (W_{ee}) from narrow (green) to broad (orange) while keeping other connections the same (gray, including W_{ee}, W_{II}, W_{Ie}) (middle panel). Cross-correlograms between E and I populations in the same column suggest that different E receptive field sizes have little impact on the spike-time correlations of local neural activity across layers (right panel).
(D) Cross-correlograms (mean ± SEM) between Narrow and 3 other cell classes in the superficial, input, and deep layer. Cross-correlations were calculated using the pooled spike trains of Narrow class and the other cell class (Board Bursty, Medium Regular, or Medium Bursty) and were averaged across sessions.
(E) The Cumming estimation plot shows the mean difference for cell-class specific comparisons of average cross-correlations between the superficial (*Super.*) and deep layers or between the input and deep layers. We picked 3 time intervals to compute the average cross-correlations (rows). The raw data of average cross-correlations is plotted on the left in each panel. Each mean difference between layers is plotted on the right as a bootstrap sampling distribution.

Supp. Figure 4 - Wang



Supp. Figure 4. Normalization Model of Attention and CCG Analyses Between Cell Classes

(A) The structure of the normalization model of attention. The left panel shows a pair of orientated grating stimuli with identical contrasts, acting as input to the model. The central black dot indicates the fixation point. The dashed red circle indicates the receptive field of the model neuron centered on the grating stimulus. The stimulus drive shown in the middle panel is a collection of neural activity driven by the stimuli. Neurons are arranged based on their receptive field center (horizontal position) and orientation preference (vertical position). The values of the stimulus drive are shown by brightness. The top panel shows the attention field as a function of the receptive field center and the orientation preference. In this case, the attention is guided to the right stimulus position and does not vary with orientation. Gray areas indicate values of 1, and white areas indicate values greater than 1. The suppressive drive at the bottom is calculated from the point-by-point product of the stimulus drive and the attention field and then pooled over space and orientation according to the suppressive field size. The stimulus drive is multiplied by the attention field and then divided by the suppressive field to generate the output firing rates of model neurons (right panel).

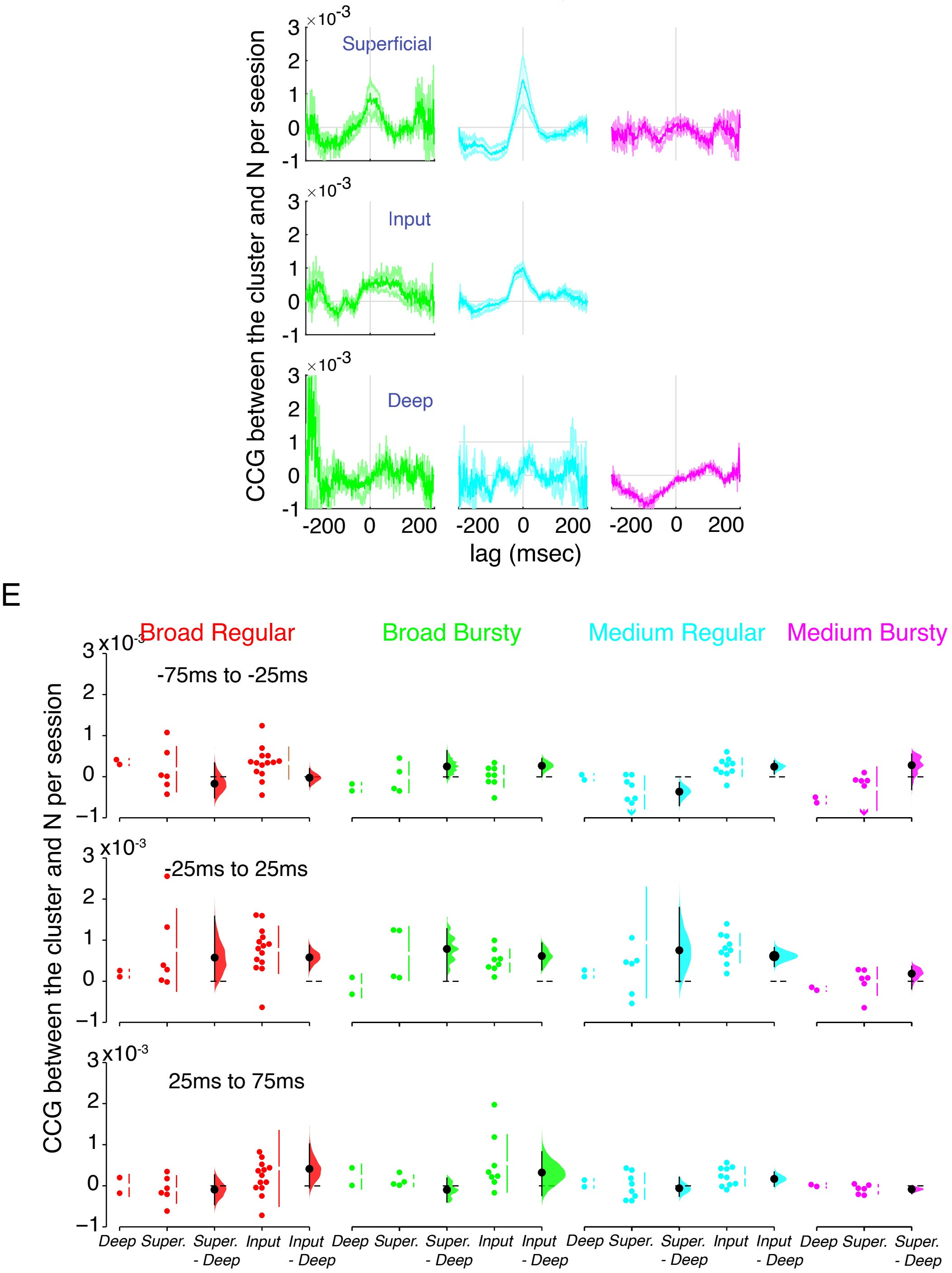
(B) i, CDIs for simulated neurons in the normalization model with different stimulus sizes and attention field sizes. In each panel, we vary the E receptive field size relative to the attention field size (x-axis), and the suppressive field size relative to the E receptive field size (y-axis). The pattern of CDI holds for a range of values of stimulus size and attention field size. ii, CDIs for simulated neurons in the normalization model with different types of inputs. We changed the stimulus drive input to the normalization model to have either a nonlinear or an attention-modulated contrast response function. We tested both the response gain (10% increase in overall response) and the contrast gain (1% of increase in detected contrast) effects. For these simulations, the attention field size is 30 and the stimulus size is 5. The pattern of CDI holds for different types of inputs.

(C) Changes in E receptive field size (white box) can also lead to the variation of CDIs across layers (left panel). We tested this hypothesis in the E-I network by adjusting the standard deviation of between-column E-E connections (W_{ee}) from narrow (green) to broad (orange) while keeping other connections the same (gray, including W_{ee}, W_{II}, W_{Ie}) (middle panel). Cross-correlograms between E and I populations in the same column suggest that different E receptive field sizes have little impact on the spike-time correlations of local neural activity across layers (right panel).

(D) Cross-correlograms (mean ± SEM) between Narrow and 3 other cell classes in the superficial, input, and deep layer. Cross-correlations were calculated using the pooled spike trains of Narrow class and the other cell class (Board Bursty, Medium Regular, or Medium Bursty) and were averaged across sessions.

(E) The Cumming estimation plot shows the mean difference for cell-class specific comparisons of average cross-correlations between the superficial (*Super.*) and deep layers or between the input and deep layers. We picked 3 time intervals to compute the average cross-correlations (rows). The raw data of average cross-correlations is plotted on the left in each panel. Each mean difference between layers is plotted on the right as a bootstrap sampling distribution.

Supp. Figure 4 - Wang



Supp. Figure 4. Normalization Model of Attention and CCG Analyses Between Cell Classes

(A) The structure of the normalization model of attention. The left panel shows a pair of orientated grating stimuli with identical contrasts, acting as input to the model. The central black dot indicates the fixation point. The dashed red circle indicates the receptive field of the model neuron centered on the grating stimulus. The stimulus drive shown in the middle panel is a collection of neural activity driven by the stimuli. Neurons are arranged based on their receptive field center (horizontal position) and orientation preference (vertical position). The values of the stimulus drive are shown by brightness. The top panel shows the attention field as a function of the receptive field center and the orientation preference. In this case, the attention is guided to the right stimulus position and does not vary with orientation. Gray areas indicate values of 1, and white areas indicate values greater than 1. The suppressive drive at the bottom is calculated from the point-by-point product of the stimulus drive and the attention field and then pooled over space and orientation according to the suppressive field size. The stimulus drive is multiplied by the attention field and then divided by the suppressive field to generate the output firing rates of model neurons (right panel).

(B) i, CDIs for simulated neurons in the normalization model with different stimulus sizes and attention field sizes. In each panel, we vary the E receptive field size relative to the attention field size (x-axis), and the suppressive field size relative to the E receptive field size (y-axis). The pattern of CDI holds for a range of values of stimulus size and attention field size. ii, CDIs for simulated neurons in the normalization model with different types of inputs. We changed the stimulus drive input to the normalization model to have either a nonlinear or an attention-modulated contrast response function. We tested both the response gain (10% increase in overall response) and the contrast gain (1% of increase in detected contrast) effects. For these simulations, the attention field size is 30 and the stimulus size is 5. The pattern of CDI holds for different types of inputs.

(C) Changes in E receptive field size (white box) can also lead to the variation of CDIs across layers (left panel). We tested this hypothesis in the E-I network by adjusting the standard deviation of between-column E-E connections (W_{ee}) from narrow (green) to broad (orange) while keeping other connections the same (gray, including W_{ee}, W_{II}, W_{Ie}) (middle panel). Cross-correlograms between E and I populations in the same column suggest that different E receptive field sizes have little impact on the spike-time correlations of local neural activity across layers (right panel).

(D) Cross-correlograms (mean ± SEM) between Narrow and 3 other cell classes in the superficial, input, and deep layer. Cross-correlations were calculated using the pooled spike trains of Narrow class and the other cell class (Board Bursty, Medium Regular, or Medium Bursty) and were averaged across sessions.

(E) The Cumming estimation plot shows the mean difference for cell-class specific comparisons of average cross-correlations between the superficial (*Super.*) and deep layers or between the input and deep layers. We picked 3 time intervals to compute the average cross-correlations (rows). The raw data of average cross-correlations is plotted on the left in each panel. Each mean difference between layers is plotted on the right as a bootstrap sampling distribution.