2023

Investigating orientation-tuned normalization in human visual cortex

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

INVESTIGATING ORIENTATION-TUNED NORMALIZATION IN HUMAN VISUAL CORTEX

by

MICHAELA KLIMOVA

BSc, University of Edinburgh, 2015
MSc, University of Edinburgh, 2016

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2023

Approved by

First Reader

Sam Ling, Ph.D.
Associate Professor of Psychological and Brain Sciences

Second Reader

Rachel Denison, Ph.D.
Assistant Professor of Psychological and Brain Sciences

Third Reader

David Somers, Ph.D.
Professor of Psychological and Brain Sciences

Fourth Reader

Melissa Kibbe, Ph.D.
Associate Professor of Psychological and Brain Sciences

ACKNOWLEDGMENTS

First of all, I would like to thank my advisor, Sam Ling, for his guidance, mentorship, and support throughout my degree. Sam is an incredible advisor with an enthusiasm for vision science, a realistic outlook, and an understanding for his students. His mentorship made my PhD experience both challenging and rewarding, and helped me to grow as a scientist over the last six years. I would also like to thank my committee members, David Somers, Rachel Denison, and Melissa Kibbe, for providing thoughtful feedback, and discussions that helped me expand my research.

I am grateful to the members of the Ling lab, Emily Wiecek, Jasmine Pan, Luis Ramirez, Anke Sun, Nick Cicero, Josh Foster, Ilona Bloem, Sara Aghajari, and Louis Vinke, for their support and friendship. The lab was always a welcoming, collaborative and fun environment, which made it so much easier to complete my PhD. Special thanks go to Ilona, Louis, and Sara, who helped me get settled in the lab and taught me essential data collection and analysis methods in the beginning of my time at BU.

Thanks are due to the Cognitive Neuroimaging Center staff; their flexibility and technical support has been instrumental in completing the imaging projects in this dissertation.

I am grateful to the BBC and GPN student community for support in classes, and for fun memories. Having time to unwind with friends was invaluable. Thank you also to my friends from home and from the UK, who have always believed in me and supported me despite the distance.

I'd like to thank my family for their lifelong support and encouragement, especially my mom for imparting her love of reading, and my dad for always encouraging me to be an independent thinker and sharing his interest in science.

Lastly, I want to especially thank my partner, Dan McKinley, for his patience, support, and encouragement throughout this journey. You were always there for me, and this would have been much more difficult – and less fun – without you.

INVESTIGATING ORIENTATION-TUNED NORMALIZATION

IN HUMAN VISUAL CORTEX

MICHAELA KLIMOVA

Boston University Graduate School of Arts and Sciences, 2023

Major Professor: Sam Ling, Associate Professor of Psychological and Brain Sciences

ABSTRACT

The brain's ability to parse sensory information into coherent scenes is underpinned by numerous mechanisms. One such mechanism is divisive normalization, a neural computation which divides neural responses to stimulation by the pooled activity of neurons responding to neighboring stimuli. Normalization strength is modulated by context; similar features (e.g., orientation) suppress each other more than dissimilar features. This feature-tuned aspect of normalization appears essential for scene segmentation, yet it has been largely understudied. Using functional magnetic resonance imaging, this work explored the feature-tuned aspect of normalization in human early visual cortex across three experiments.

Experiment 1 examined the degree to which suppression in visual cortex is sensitive to orientation differences, by presenting observers (n = 10) with a bandpass-filtered noise stimulus and parametrically varying the orientation difference between its components. Consistent with prior work, parallel stimuli caused strongest BOLD signal suppression, while orthogonal stimuli caused least suppression. Extending previous findings, I obtained a measure of the orientation-tuned suppression function across a

spectrum of orientation differences and estimated its bandwidth, providing a link between orientation difference and suppression strength.

Experiment 2 tested the effects of visual attention on orientation-tuned suppression. I examined whether attention alters its bandwidth by measuring the orientation-tuned suppression function of 10 observers while they attended either to the oriented stimulus or away from it. Attention increased the BOLD signal magnitude across all orientation differences (consistent with prior findings), but did not change other aspects of the function. Therefore, the specificity of orientation-tuned normalization appears unaltered by attention.

Finally, Experiment 3 explored the orientation tuning of normalization by measuring contrast response functions (CRF) to a central oriented stimulus suppressed by either a parallel or orthogonal high-contrast annulus (n = 10), and quantifying the resulting CRF with a normalization model. Consistent with non-human animal research, the relative orientation of center and surround impacted the CRF, with signs of stronger suppression seen with parallel surround. Interestingly, these modulations were spatially specific and largely limited to measurements from the boundary between the central and surrounding stimulus. Together, these results help extend our understanding of contextual effects in the human visual cortex.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
BOLD	Blood Oxygen Level-Dependent
cpd	cycles per degree
CRF	Contrast Response Function
dva	Degree Visual Angle
EPI	Echo-Planar Imaging
fMRI	Functional Magnetic Resonance Imaging
FoV	Field of View
FS-FAST	FreeSurfer Functional Analysis Stream
FWHM	Full Width Half Maximum
GLM	General Linear Model
Hz	Hertz
mm	Millimeter
MPRAGE	Magnetization-Prepared Rapid Acquisition with Gradient Echo
ms	Millisecond
MT	Middle Temporal Cortex
pRF	Population Receptive Field
PSC	Percent Signal Change
RF	Receptive Field
RM	Repeated Measures
ROI	Region of Interest

RSVP	Rapid Serial Visual Presentation
s	Second
SD	Standard Deviation
SEM	Standard Error of the Mean
T	Tesla
TE	Echo Time
TR	Relaxation Time
V1	Primary Visual Cortex
V2	Secondary Visual Cortex
V3	Third Visual Cortex

CHAPTER ONE: BACKGROUND

Preamble

When navigating our environment, the visual system relies heavily on contextual information to help us rapidly and efficiently parse scenes into meaningful units. A classic example is the pop-out effect, a perceptual phenomenon where the detection of a visual stimulus (such as a visual search target) is rapid and effortless when the appearance of the target object contrasts sharply with its surroundings in basic visual features such as color, while similar or identical features between the target and its surroundings hinder the speed and efficiency of visual search (Toth et al., 1996; Treisman & Gelade, 1980). Contextual effects play a role in essential aspects of scene processing such as grouping (Schwartz et al., 2009), contour integration (Kapadia et al., 1995), or object detection and figure-ground segregation through texture discontinuities (Schwartz & Simoncelli, 2001). A large body of scientific literature indeed suggests that the visual system makes use of spatial context to achieve efficient representation of visual information (Coen-Cagli et al., 2015; Schwartz & Simoncelli, 2001; Simoncelli & Olshausen, 2001).

The computations responsible for contextual effects begin to operate in the early stages of the visual system, as a result of interactions among visual neurons. The response of a visual neuron is modulated, typically suppressed, by placing additional stimuli in the vicinity of its classical receptive field (RF) (Cavanaugh et al., 2002a; DeAngelis et al., 1994; Polat et al., 1998; Sengpiel et al., 1998; Webb et al., 2005). The presence and nature of suppression is determined by the relative properties of the central RF and the

additional, surround stimuli; in particular, stimuli similar in features such as orientation suppress each other to a greater extent than dissimilar stimuli (Cavanaugh et al., 2002a; Self et al., 2014; Webb et al., 2005). These observations are consistent with efficient coding theories, which propose that the visual system attempts to reduce redundancy contained in visual scenes (i.e., make neural responses as independent of each other as possible (Carandini & Heeger, 2012; Schwartz & Simoncelli, 2001; Simoncelli & Olshausen, 2001)), for example by more strongly suppressing similar information coming from adjacent areas in a scene, and diminishing suppression when adjacent image areas differ, as they are likely to signal important information such as the presence of a figure-ground boundary (Geisler et al., 2001; Schwartz & Simoncelli, 2001).

Many of these properties of early visual responses are captured well by a computational principle termed *divisive normalization*, which dictates that the neural response to a stimulus is regulated by a pool of responses of neighboring neurons (Carandini et al., 1997; Carandini & Heeger, 2012; Heeger, 1992). While divisive normalization has been used to explain a wide range of visuocortical phenomena, as well as ways in which sensory neural responses are affected by cognitive influences such as attention (Reynolds & Heeger, 2009), the feature-tuned aspect and its likely importance in shaping visual perception is underemphasized in current human neuroimaging research. The goal of this body of work is to more closely examine and characterize the properties of feature-tuned normalization in the human visual cortex, and to explore interactions between this effect and other cognitive and perceptual processes.

Contextual effects in early visual system

The visual system is highly structured. Bottom-up sensory input flows in a hierarchical pathway from the retina of the eye, to the lateral geniculate nucleus (LGN) of the thalamus in subcortex, and from there to the primary visual cortex (V1), extrastriate areas (secondary visual cortex; V2, third visual cortex; V3), and higher-order visual areas. In V1, the earliest cortical visual area, neurons are highly selective for basic stimulus features such as orientation, motion direction, or spatial frequency (De Valois et al., 1982; Hubel & Wiesel, 1968), such that their firing rate is highest when this feature is found in their receptive field (RF; the portion of the visual field where visual stimulation elicits a response from the cell), and they show baseline or below-baseline activity when non-preferred features are present instead. In addition to feature optimality, the intensity (spatial contrast) of a stimulus also determines the strength of responses; a typical V1 cell will increase its firing with increasing stimulus contrast, with a plateauing, or saturation, of this function at the highest contrasts; the resulting S-shaped curve describes a representative contrast response function (CRF) (Albrecht & Hamilton, 1982). As one traverses the visual hierarchy from V1 to V2, V3, and higher-order visual areas, visuocortical RFs considerably increase in size (Gattass et al., 1988; Kastner et al., 2001; Ungerleider et al., 1984). It is thought that neurons in each successive stage integrate feedforward inputs across a number of cells from the lower visual areas in order for the cells to integrate the basic features detected by earlier areas across space and assemble them into more complex representations (DiCarlo et al., 2012; Pasupathy & Connor, 2002). In higher-order visual areas, such as the inferotemporal cortex, neurons have much

larger RF sizes compared to early visual cortex, and are specialized in detecting increasingly complex features such as entire object contours, which can span large areas of the visual field (Gross et al., 1969). In addition to feedforward neural pathways feeding the visual signal in various stages of processing along the hierarchy, there are also feedback connections, from higher areas back to lower areas; these are thought to have some role in regulating the firing of cells in lower areas, as they can provide contextual information about the broader visual scene (Angelucci et al., 2017; Angelucci & Bressloff, 2006).

Crucially, within each visual area, cells are retinotopically organized, meaning that neurons which are positioned near each other on the cortical surface receive stimulation from adjacent portions of the visual field. Interactions between RFs within each area are made possible through anatomical connections linking neurons both in close proximity and over larger distances on the cortical surface, facilitating surround modulation from areas immediately adjacent to the RF, but also over larger cortical distances (Angelucci et al., 2017; Angelucci & Bressloff, 2006; Bair et al., 2003). The anatomical connections and structure described above lends itself to extensive neural interactions which allow for the integration of visual information across different locations of the visual field, a process necessary for a unified percept of the broad visual scene.

Nature of surround modulation

In early electrophysiology experiments with cats and primates aimed at mapping the extent of the V1 cells' RF with simple grating stimuli, a number of research groups

noted that the cell response magnitude depends on stimulus size. Increasing the size of the stimulus led to increasing firing rates, up until the stimulus reached a certain size, after which the response peaked and subsequently decreased with increasing stimulus size (DeAngelis et al., 1992). The area across which the stimulus elicits increasing responses is thought of as the excitatory RF, also known as summation field (Cavanaugh et al., 2002a; DeAngelis et al., 1992), while the area beyond it is termed the RF surround. The properties of surround modulation have since been widely studied. In sum, the response of a typical V1 neuron is modulated by placing other stimuli outside of the excitatory RF, even though these surround stimuli evoke no spiking response when presented on their own (Carandini et al., 1997; Cavanaugh et al., 2002a; DeAngelis et al., 1994; Sengpiel et al., 1998; Webb et al., 2003, 2005). The effect of adding a surround stimulus is usually a decrease in activity of the center RF (surround suppression), especially with high-contrast surround stimulation; however, in some configurations, the addition of an external stimulus can instead increase the cell response (surround facilitation) (Knierim & Van Essen, 1992; Levitt & Lund, 1997; Toth et al., 1996).

The relative properties of the center and surround are crucial in determining the effects of the surround. Specifically, the amount of suppression depends on feature similarity: the prevalent finding is that surround stimuli which share similar or identical features with the center, such as motion direction (Cavanaugh et al., 2002b; Levitt & Lund, 1997), spatial frequency (Webb et al., 2005), and orientation (Self et al., 2014; Trott & Born, 2015; Webb et al., 2005), elicit stronger suppression of neural activity than dissimilar stimuli (e.g. orthogonal orientation of surround with respect to center).

However, in some instances, the addition of a collinear surround can lead to facilitation of the neural response, as opposed to suppression; this has been reported especially when the center RF contrast is low (Ichida et al., 2007; Polat et al., 1998; Toth et al., 1996), when there is a gap between the center and surround stimuli (Henry et al., 2013) or when the center and surround are otherwise suboptimally stimulated (Henry et al., 2013; Polat et al., 1998; Shushruth et al., 2012). Orthogonally oriented surrounds can also sometimes cause facilitation under these conditions (Shushruth et al., 2012; Sillito et al., 1995). However, not all center-surround experiments report surround facilitation (Cavanaugh et al., 2002a; DeAngelis et al., 1994; Sengpiel et al., 1998). Beside the optimality of the RF stimulus, another likely cause of differences between studies is the presence or absence of a gap between the center and surround gratings (Shushruth et al., 2012), which may change whether the surrounding stimulus also directly activates the center RF and its LGN afferents (Ichida et al., 2007; Shushruth et al., 2012). Generally, surround facilitation is less well studied than suppression effects, which are more commonly found.

Surround suppression does not act uniformly across time or cortical space. In the temporal domain, research suggests that there are two components: one has a rapid onset and is untuned to stimulus properties (i.e., no difference in suppression strength regardless of the relative orientation of center and surround), while the second component is delayed and more sharply tuned (Henry et al., 2013; Self et al., 2014; Webb et al., 2005). Moreover, surround modulation also differs depending on the distance between the center and the surround (Angelucci et al., 2017; Shushruth et al., 2013); in the

surround location immediately near the center, suppression appears to be stronger and more sharply tuned compared to the far surround region, which extends far beyond the RF (some studies find suppression arising from surround regions as large as 12.5-13° in diameter (Sceniak et al., 2001; Shushruth et al., 2009)). These properties have led to the proposal that multiple mechanisms, and different sets of connections may be involved in bringing about suppression near the RF center than those responsible for far surround modulation (Angelucci et al., 2017; Angelucci & Bressloff, 2006). Specifically, the fast and untuned (feature-agnostic) component of suppression has been theorized to rely on feedforward connections from the LGN, whose cells also exhibit some untuned surround suppression. The near-surround, sharply orientation-tuned component could be generated by horizontal connections, which link neurons with the same orientation preference (Angelucci & Bressloff, 2006). The conduction velocity is slower for horizontal connections, which would also explain the slower onset of the tuned component. The modulation coming from far surround can be explained by the involvement of feedback connections, as within-area connections do not span sufficient distances to account for the spatial extent of far-surround modulation (Angelucci et al., 2017).

Evidence from human psychophysics and neuroimaging

Beyond animal electrophysiology, surround suppression is readily replicated in human studies. In psychophysical investigations, adding a high-contrast surround to a central grating most often decreases subjective center contrast and raises perceptual thresholds for the center stimulus (Xing & Heeger, 2000, 2001). In functional neuroimaging such as with BOLD fMRI, surround stimuli often cause a reduction in the

BOLD signal, considered an indirect measure of the neural response. Generally, there is good agreement between stimulus configurations that cause perceptual and neural suppression (Xing & Heeger, 2001). The sensitivity to feature differences observed in animal results is also found in human research; perceptual suppression is stronger in center-surround stimuli with identical features such as motion direction (Phillips et al., 2021) and orientation (Cannon & Fullenkamp, 1991; Petrov et al., 2005; Solomon et al., 1993). Some studies also find surround facilitation in configurations where center stimulus contrast is considerably greater than that of the surround (Cannon & Fullenkamp, 1993; Xing & Heeger, 2001). Differences in suppression strength, and its sensitivity to feature similarity, exist between foveal and peripheral vision; perceptual suppression is stronger in the periphery (including with a low-contrast surround) and is not feature-tuned, while near the fovea, suppression is tuned and much weaker when the center and surround differ in orientation by more than 45 degrees (Xing & Heeger, 2000). In occipital EEG, which enables the measurement of visual-evoked potentials with precise timing, findings support the existence of two surround modulation components; an earlier visual response component to a central grating was equally suppressed by orthogonal or parallel surrounds, while a later component was more strongly suppressed by a parallel surround (Schallmo et al., 2019). In fMRI, center-surround displays evoke suppression of the BOLD signal, which is stronger when the surrounding stimulus is collinear with the center, as opposed to orthogonal (McDonald et al., 2009; Pihlaja et al., 2008; Poltoratski et al., 2017; Williams et al., 2003; Zenger-Landolt & Heeger, 2003). While some animal and psychophysical studies tested a range of orientation offsets

between the center RF and the suppressive annulus (Cannon & Fullenkamp, 1991; Phillips et al., 2021; Trott & Born, 2015), human fMRI studies have mostly focused on contrasting collinear and orthogonal surrounds (Chen, 2014; Joo et al., 2012; Ling et al., 2015; McDonald et al., 2009; Pihlaja et al., 2008; Poltoratski et al., 2017). In general, human neuroimaging experiments have not interrogated surround suppression properties at the same level of detail as electrophysiological studies.

Relation to natural scene processing

Surround modulation, and its feature tuning property, are thought to serve an important function in natural scene processing. Natural scenes are complex and the objects, textures, and contours within them usually extend beyond any single visual cell receptive field (particularly in earlier visual areas such as V1, where RFs are smaller compared to higher-order areas), making surround modulation a ubiquitous occurrence across visual cortex. The contextual effects described in the previous section are theorized to translate into differences in subjective salience across the visual scene, which aid in figure-ground segregation or target detection (Li, 2002). For example, stronger suppression from a collinear surround results in weaker neural signal, thereby discounting inputs coming from regions of a scene that carry identical information to neighboring regions. Conversely, weak or no suppression as a result of orthogonal features would translate into stronger perceptual salience in regions containing discontinuities of texture, orientation, or contrast, thus enhancing the perceived salience of object borders and contours (Coen-Cagli et al., 2012; Schwartz & Simoncelli, 2001).

Furthermore, research strongly suggests a link between the nature of feature sensitivity of suppression and natural scene statistics. It has long been theorized that visual neurons are adapted and optimized to the predominant statistical biases and regularities in the environment, such that they are most efficient at processing features that occur more frequently (Simoncelli & Olshausen, 2001). A typical visual scene contains a lot of redundant information, as nearby areas are most likely to be very similar in texture, orientation, and intensity, and thus the information they contain is highly correlated with nearby locations (Schwartz & Simoncelli, 2001). Most natural scenes also carry a signature power spectrum, where the spectral power falls with spatial frequency (Hibbard & O'Hare, 2015; Simoncelli & Olshausen, 2001). According to theories of efficient coding, the visual system tries to reduce correlations between outputs of individual neurons, to achieve sparse coding (Atick, 1992). The above-mentioned environmental regularities are thought to be exploited by the visual system to achieve this. For instance, in natural images, nearby edges tend to contain the same orientation (Geisler et al., 2001), and the similarity of features between two points in a visual scene decreases with their distance (Schwartz & Simoncelli, 2001). Therefore, discounting (through surround suppression) signals from adjacent areas which contain similar information would reduce the amount of redundant information carried by nearby neurons (Olshausen & Field, 1997). In support of efficient coding, it has been shown that the responses of V1 neuron pairs become decorrelated and that coding sparseness is increased when both the RF and surround are stimulated in natural scene viewing (Vinje & Gallant, 2000).

Divisive normalization model

Surround suppression, as well as other typical, nonlinear behaviors of early visual neurons such as contrast saturation, are thought to serve as a type of sensory gain control, solving the issue of limited dynamic range and maintaining response sensitivity (Carandini et al., 1997; Ohzawa et al., 1985). Experiments combining electrophysiology and computational modeling showed that many such nonlinearities can be explained by adding a normalization (division) stage to the initially linear response of V1 neurons (Carandini et al., 1997; Heeger, 1992); this observation has been formalized in the divisive normalization model. Under divisive normalization, the response magnitude of a visual neuron is dependent on its own activity in response to a stimulus, as well as on the activities of a pool of neighboring neurons (the normalization pool), in that it is divided, or normalized, by the weighted combined responses of itself plus the summed activity of the normalization pool (Carandini et al., 1997; Carandini & Heeger, 2012). Therefore, divisive normalization mathematically expresses the idea that the response of a single unit depends not just on the stimulus in its RF, but also on the spatial context of that stimulus, which comes from beyond the RF. The model has proven powerful in accounting for early sensory processing phenomena as well as some higher-order perceptual and cognitive processes (Carandini & Heeger, 2012; Louie et al., 2011; Rabinowitz et al., 2011). There is also evidence that deficient normalization of neural responses can contribute to perceptual abnormalities seen in disease. For instance, patients with schizophrenia show impaired contour detection (Schallmo et al., 2013), or weaker center-surround suppression effects (Pokorny et al., 2023). In photosensitive

epilepsy, contrast-response functions measured with electroencephalography (EEG) show a lack of saturation at high contrast, compared to those of healthy controls (Porciatti et al., 2000).

Feature tuning can be to some extent incorporated into the divisive normalization model to account for differences in suppression under different configurations (Coen-Cagli et al., 2015). In the context of natural scenes, a more flexible model where normalization is engaged when the RF center and surround are similar in basic features and "switched off" when they differ, has been shown to better match the responses of V1 neurons shown images of natural scenes, compared to the default normalization model in which the normalization step always takes place (Coen-Cagli et al., 2015). These results again highlight the significance of feature-tuning in the modulation of neural responses by spatial context. However, it is unclear whether the divisive normalization model can account for surround facilitation that is sometimes seen. One fMRI and psychophysical study suggested that both can be explained by the normalization model; facilitation could be obtained if the suppression (denominator) term in the model is very low, as would be the case with a small, low-contrast central stimulus (Schallmo et al., 2018). However, this study tested motion direction perception, not orientation, and found facilitation only in later visual areas, and not the early visual cortex (V1).

Attentional effects in sensory gain control

As discussed above, gain control mechanisms such as surround suppression are thought to improve visual scene processing due to increasing the saliency of relevant scene areas, such as figure-ground boundaries. Endogenous attention is a top-down

process which likewise contributes to faster, more efficient scene processing and object detection. It is widely thought to do so by operating on the same neural circuits that mediate gain control (Raizada & Grossberg, 2003; Reynolds & Chelazzi, 2004). Attending to a particular area of a scene (spatial attention) or visual feature (feature-based attention) increases the subjective salience of the attended objects (Carrasco et al., 2004), a finding corroborated by enhanced task performance (Carrasco, 2011). At the neural level, attention causes profound gain changes, and visual responses are subject to attentional enhancement (Desimone & Duncan, 1995; McAdams & Maunsell, 1999), and these changes are largely multiplicative (Martinez-Trujillo & Treue, 2004; McAdams & Maunsell, 1999), scaling responses in a way that contributes to decreased contrast threshold and better perceptual discrimination and sensitivity (Reynolds et al., 2000; Reynolds & Chelazzi, 2004). In addition to facilitating the processing of single stimuli, attention also enables rapid selection of stimuli relevant to the current goal or task (Moran & Desimone, 1985). Electrophysiological experiments in areas V4 and MT showed that multiple stimuli are presented together in one RF, for example one having the nonpreferred motion direction and the other with preferred motion direction, neural responses increase if the preferred stimulus is attended, but decrease if the nonpreferred stimulus is attended (Treue & Trujillo, 1999). Similarly, when attention is directed to the RF surround stimulus, the RF response also decreases (Motter, 1993). Therefore, the attended stimulus preferentially determines the neural response. The observed multiplicative changes suggest that attention acts in a way so as to mimic increasing the effective attended stimulus contrast.

One proposed mechanism of this attentional effect on visuocortical responses is that attention acts to reduce suppressive contextual effects, effectively releasing the neural response from divisive normalization. This process is formally described in the normalization model of attention, which can account for a number of electrophysiological findings, including enhancement vs. reduction of neural responses, depending on the stimulus and the spatial and featural extent of the attentional window (Reynolds & Heeger, 2009). Links between attention and normalization are supported by experimental data; in electrophysiological studies, the strength of attentional modulation of individual neurons in area MT was found to correlate with the strength of normalization (Lee & Maunsell, 2009; Ni & Maunsell, 2017).

In neuroimaging research, attention is consistently found to increase BOLD responses (Buracas & Boynton, 2007; Itthipuripat et al., 2019; Murray, 2008; Somers et al., 1999) and EEG signal (Itthipuripat et al., 2019). However, it has been challenging to study the nature of attentional modulation of BOLD responses; unlike in electrophysiological studies, BOLD responses appear to increase additively, although some multiplicative changes have been observed, particularly in feature-based attention (Foster & Ling, 2022). Some fMRI evidence also supports the relationship between normalization and attention discussed above; early visual voxels which exhibit stronger orientation-tuned normalization also show stronger attentional modulation (Bloem & Ling, 2019).

Organization of dissertation

This body of work aims to more closely examine and characterize the properties of orientation-tuned normalization in the human visual cortex. The three main chapters in this dissertation contain three fMRI experiments addressing different aspects of how feature-tuned normalization operates in the early human visual cortex and how it interacts with other basic visual properties, as well as top-down processes. In the next chapter, I first explore the orientation-tuned property of normalization by measuring human visual cortex BOLD responses to full-field fixed-contrast stimuli with parametrically varied orientation differences, and reconstruct a function describing the relationship between suppression strength and stimulus orientation difference. In the third chapter, I modify this paradigm to investigate how orientation-difference sensitivity is affected by directing endogenous visual attention to the stimulus, as opposed to passive observation. Finally, in the third chapter, the focus is on the interplay between orientation difference tuning and the role of contrast in cortical responses; here I measure the full contrast response across early visual cortex to a central grating stimulus under suppression by either collinear or orthogonal surround stimulus, and characterize the impact of tuned suppression on the population contrast response function.

CHAPTER TWO: THE SPECIFICITY OF ORIENTATION-TUNED NORMALIZATION WITHIN HUMAN EARLY VISUAL CORTEX[1]

Introduction

Our visual system is tasked with representing our environment as completely as possible while incurring the lowest possible metabolic cost. To accomplish this, the brain relies on gain control mechanisms, which are believed to play an essential role in reducing redundancy in neural coding (Carandini & Heeger, 2012; Coen-Cagli et al., 2015; Schwartz & Simoncelli, 2001; Simoncelli & Olshausen, 2001). One of the signature examples of gain control in action is surround suppression, wherein the response of a visual neuron to a stimulus is divisively normalized by the sum of activity generated by the stimulus and that generated by a pool of neighboring neurons (Carandini et al., 1997). Divisive normalization not only demonstrably shapes visual responses (Carandini et al., 1997; Cavanaugh et al., 2002a; Flevaris & Murray, 2015; Heeger, 1992; Ni et al., 2012; Self et al., 2014; Shushruth et al., 2013; Tsai et al., 2012; Xing & Heeger, 2000), but also plays a role in other sensory systems (Rabinowitz et al., 2011) and higher-level cognitive processes (Louie et al., 2011), and has become regarded as a canonical neural computation (Carandini & Heeger, 2012).

Within visual cortex, the strength of normalization is modulated by stimulus features, and depends on the degree of feature similarity (Tsai et al., 2012). Indeed, normalization-driven suppression has been shown to be strongest when stimuli in the

[1] This chapter is based in full on the previously published article: Klímová, M., Bloem, I. M., & Ling, S. (2021). The specificity of orientation-tuned normalization within human early visual cortex. Journal of Neurophysiology, 126(5), 1536-1546. https://doi.org/10.1152/jn.00203.2021

inhibitory surround share similar orientation content to that of stimuli in the excitatory center (Bloem et al., 2018; Bloem & Ling, 2019; Petrov et al., 2005; Schallmo & Murray, 2016; Trott & Born, 2015; Tsai et al., 2012; Xing & Heeger, 2000). This unique orientation-tuned property of normalization has been proposed to contribute to the efficient neural coding of figure-ground segmentation in visual scenes. In natural scenes, feature discontinuities tend to indicate borders between objects, while regions with high feature similarity are often close to each other, and belong to the same object (Angelucci et al., 2017; Schwartz & Simoncelli, 2001). Orientation-tuned normalization is believed to play a key supporting role in the efficient coding of neural representations, compressing the amount of neural resources coding redundant collinearities and instead dedicating more vigorous neural responses to discontinuities in orientation, which typically signify areas of visual salience, such as the border between figure and ground (Coen-Cagli et al., 2012; Geisler et al., 2001).

Despite evidence supporting orientation-tuned normalization in the visual cortex of animals (Self et al., 2014; Trott & Born, 2015; Webb et al., 2005), investigations in humans mostly focus on contrasting responses to just two cases at the extremes of orientation similarity: collinear and orthogonal stimuli (Bloem & Ling, 2019; Chen, 2014; Flevaris & Murray, 2015; Joo et al., 2012; Ling et al., 2015; McDonald et al., 2009; Pihlaja et al., 2008; Poltoratski et al., 2017; Williams et al., 2003). In fMRI, stronger suppression for collinearly configured stimuli, compared to orthogonal configurations, has been demonstrated using center-surround grating displays (McDonald et al., 2009; Pihlaja et al., 2008; Williams et al., 2003), flanker displays (Chen, 2014;

Flevaris & Murray, 2015; Joo et al., 2012; Poltoratski et al., 2017), and overlapping gratings (Bloem & Ling, 2019; Ling et al., 2015), establishing that BOLD signal suppression from surrounding stimuli is context-dependent. However, such comparisons do not allow for closer characterization of this feature-tuned aspect of normalization, for which one would need to parametrically manipulate stimulus orientation differences in finer steps. Capturing the bandwidth, or specificity, of tuned normalization could improve existing models of divisive normalization by estimating the degree of feature similarity required to engage suppression beyond a simple collinear-orthogonal distinction. To our knowledge, no human neuroimaging studies have characterized the specificity of orientation feature-tuning. A narrow bandwidth, or high specificity, would suggest that only a small deviation in orientation similarity is sufficient to decrease suppression strength to produce a noticeable change in the BOLD response. On the other hand, a broad bandwidth would suggest a large change in orientation similarity is needed in order to modulate suppression strength and BOLD responses. The goal of our experiment was to provide a thorough investigation of the specificity of these feature-selective properties in humans.

In this study, we used functional magnetic resonance imaging (fMRI) to characterize the orientation bandwidth of tuned normalization. In order to measure orientation-tuned normalization on a large population scale, rather than use a typical center-surround configuration, we designed full-field circular stimuli consisting of wedge-shaped components containing orientation information, and manipulated the orientation differences between the neighboring stimulus components. The extent to

which the differently oriented components suppress each other depended on the bandwidth of orientation-tuned suppression. Our results revealed a gradual decrease in BOLD response as orientation similarity increases. We characterized the bandwidth within and across early visual areas V1-V3 by fitting Gaussian functions to the data at the visual area level, as well as at the level of individual voxels. Interestingly, the estimate of cortical bandwidths for tuned suppression dovetails nicely with known natural scene statistics (Sigman et al., 2001), supporting efficient coding approaches according to which the visual system is adapted to natural scene statistics (Simoncelli & Olshausen, 2001; Vinje & Gallant, 2000).

Materials and Methods

Observers

The study was approved by the Boston University Institutional Review Board. 10 observers (6 female) participated in the experiment, each completing two scan sessions. An additional observer took part but was excluded from the study after the first session due to poor behavioral performance (fixation task accuracy <75%; see below). All observers reported normal or corrected-to-normal vision. All were between the ages of 18 and 40, provided written informed consent and received monetary compensation for their participation, with the exception of two observers who were the authors of the study.

Apparatus and Stimuli

Stimuli were generated using the Psychophysics toolbox in MATLAB (R2015b) on a MacBook Pro (OS X version 10.10.5), and displayed on a rear-projection screen with a gamma-corrected projector (ProPixx DLP LED, VPixx Technologies). The circular stimulus had an outer diameter subtending 17 degrees visual angle (dva) and inner diameter 3 dva. At the innermost display eccentricity (1.5 dva from central fixation), each wedge subtended roughly 0.47 dva and at the outermost eccentricity each wedge subtended 2.66 dva. Stimuli were generated by orientation and spatial frequency band-pass filtering white noise, keeping only spatial frequencies between 2 and 3 cycles per degree and the specified orientations, with an orientation filter bandwidth of 10º. Stimuli were presented at 50% Michelson contrast, on a mean luminance background.

We generated stimuli containing 20 wedges in two interleaved sets of 10, with 7 possible equally spaced orientation differences between the two sets (0º, 15º, 30º, 45º, 60º, 75º and 90º; **Figure 2.1**). In order to avoid strictly orientation-specific effects in our data, we used 6 different "base" orientations (0º, 18º, 36º, 54º, 72º, 90º) – these orientations also track the angular values of the individual wedges. All stimuli were generated starting from 0º orientation, which we designated as vertical; a *change in base orientation* means that the entire stimulus was rotated by the new base orientation value. For instance, if the orientation difference was 30º and the base orientation was 18º in a given condition, the actual orientation content of the two sets of wedges would be 18º and 48º, maintaining the 30º orientation difference but varying the starting orientation. The orientation content of the wedges (which orientation is displayed within which set of 10) was also counterbalanced; for example, if the orientation difference was 15º and the base

orientation was 0º, this condition would be presented four times during each session such that the sets of 10 wedges containing each orientation were switched twice between the four blocks of that specific condition (combination of orientation difference and base orientation). The white noise patches within each half of the stimulus were independent of each other.

Procedure

Data were acquired at the Cognitive Neuroimaging Center at Boston University using a 3.0 Tesla Prisma scanner (Siemens, Erlagen, Germany) with a 64-channel head coil. All participants underwent two scanning sessions, each lasting between 2-2.5h. In each session, observers completed 14 functional scan runs (with the exception of one who only completed 11 runs in one of their sessions) with the acquisition field of view oriented perpendicular to the calcarine sulcus (T2*-weighted in-plane multi-slice imaging sequence with multiband factor 3 (Moeller et al., 2010; Xu et al., 2013), 36 slices, TR = 1s, TE = 35.4ms, flip angle = 64º, FoV = 136mm, voxel size = 2mm isotropic). An anatomical scan used to register the functional data was acquired during a separate session (T1-weighted multi-echo MPRAGE 3D sequence, voxel size = 1mm isotropic, FoV = 256x256x176mm, TR = 2,530ms, TE = 1.69ms, flip angle = 7º).

Stimuli were presented in a block design (14 seconds on, 14 seconds off). Each functional run lasted 350 s (1s TR) and contained 12 stimulus blocks. Each set of 14 runs contained 24 repetitions of each orientation difference condition, fully counterbalanced to reach equal numbers of base orientation presentations. The noise patch underlying the orientation stimulus was refreshed at 2 Hz, while within each stimulus block of the run,

the orientation condition remained unchanged. Throughout the run, participants were required to maintain central fixation and perform a letter discrimination task within the fixation circle. The fixation circle subtended 0.8 dva and contained a stream of letters, refreshing every 250 ms. Participants' task was to detect letters "J" and "K" within this distractor letter stream and press a corresponding key on a two-button response box for each target letter as soon as it was detected. The letter detection/discrimination task was employed throughout the run in both baseline and stimulus blocks, and targets could appear at any point during the run (regardless of baseline or stimulus block). Behavioral performance in the fixation letter detection task averaged 93.2% accuracy (±1.86% SEM), confirming that our observers were well able to maintain alert fixation throughout the scan session. Prior to acquiring the task scans, participants underwent two runs of a functional localizer (14 s on, 14 s off, 182 TRs total), later used for voxel selection. The localizer stimulus was composed of a 100% Michelson contrast spatial pattern flickering at a rate of 10 Hz. The localizer stimulus had identical inner and outer eccentricity bounds to the experimental stimulus and encompassed the entire stimulus display (rather than split into individual wedges). The spatial pattern of the localizer was created by summing and rectifying pairs of spiral and radial gratings. Throughout localizer presentation, participants were engaged in a fixation task identical to the main experiment (see above). In both the main task and localizer, each scan always began and ended with an off-block.

fMRI data analysis

Population receptive field mapping and voxel selection. Population receptive field (pRF) mapping was conducted for each participant in a separate session, using the analyzePRF toolbox for MATLAB (Kay et al., 2013; Kriegeskorte et al., 2008). In the pRF session, observers were presented with 2-3 scans of each of two types of mapping run: expanding/contracting ring and bar sweep stimuli, and rotating wedge stimuli. The stimuli were composed of a pink noise background with color objects and faces of varying spatial scale, on a display with mean luminance background (Benson et al., 2018; Kriegeskorte et al., 2008). These are the same stimuli as the ones used in the Human Connectome Project 7T Retinotopy dataset (Benson et al., 2018). The pink noise and images were refreshed at a rate of 15 Hz. The results were analyzed with analyzePRF, implementing the compressive spatial summation pRF model (Kay et al., 2013). The results of the pRF analysis were used to manually draw region-of-interest (ROI) labels, defining early visual areas V1, V2 and V3. pRF modeling results were then used in conjunction with the localizer data to select voxels for further analysis. Within each label, we first identified the top 40% most visually responsive voxels across both sessions, based on the localizer data. From this subset, we further excluded voxels whose population receptive fields were located outside the eccentricity bounds of the stimulus, or those whose pRF model fit was poor ($R^2 < 10\%$). On average, this procedure left 446 ± 96 (SD) voxels in V1, 394 ± 62 voxels in V2, and 281 ± 45 voxels in V3 for analysis.

fMRI data preprocessing. We applied EPI distortion correction to all fMRI BOLD time series data using a reverse-phase encoding method (Andersson et al., 2003)

implemented in FSL (Smith et al., 2004). The field map corrected data from each session were then preprocessed with standard motion correction procedures, Siemens slice timing correction, and boundary-based registration between functional and anatomical spaces (Greve & Fischl, 2009) implemented in FreeSurfer (Fischl, 2012), version 5.3. No spatial smoothing was applied. To achieve voxel-by-voxel alignment within and between the two experimental sessions, we applied robust rigid registration (Reuter et al., 2010), using the middle time point of each run, and aligned each functional run from both sessions to the first localizer scan in the first session, which had been aligned to the anatomical data in the boundary-based registration step. Data for each voxel were subsequently detrended, high-pass filtered and converted to percent signal change using custom MATLAB scripts. The functional localizer data were analyzed separately for each session using a standard GLM analysis in FreeSurfer. The main task data for the two sessions were concatenated and further analyses were performed in MATLAB using custom code.

fMRI data analysis. For each observer, data were summarized in every voxel by computing an event-triggered average for each orientation difference condition, collapsing across the 6 base orientations and across repetitions of each condition. To account for hemodynamic delay, a temporal shift of 6 TRs (6 s) was implemented prior to averaging, resulting in an averaging window starting at 6 TRs post stimulus onset and ending at 14 TRs (stimulus offset), which captured the peak of the BOLD activity.

Model fitting. We quantified the tuned normalization parameters in two ways: on the ROI-averaged data, and in individual voxels. To obtain the ROI mean, we averaged across all voxels within each ROI for each observer, and performed the fitting procedure

described below for each ROI and observer. In individual voxels, model fitting was done on each voxel's averaged response. To quantify the tuned normalization parameters, the data were fit with a half Gaussian function centered on an orientation difference of 0º:

$$Response = b + Ae^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (1),$$

where *b* refers to the baseline BOLD activation, *A* is the amplitude or the peak of the function, μ is the mean of the function (set to 0º in order to constrain the fit), and σ is the standard deviation, which we will use as an estimate of bandwidth. For each voxel or ROI, the Gaussian was fit to the 7 BOLD response values for each orientation difference by minimizing least-squared error using the MATLAB *fmincon* function. The bandwidth parameter was constrained between 5 and 90 degrees; this was done to avoid distortions caused by voxels that either showed no tuning to orientation difference (resulting in a flat response function) or unreasonably narrow tuning. Furthermore, due to the orientation filter bandwidth of 10º (5º on either side of the target orientation), the actual minimum possible angular difference between individual orientation segments in two components was 5º (in the 15° orientation difference condition). The overall response amplitude was constrained by an upper bound of 10% signal change; no constraints were applied to the baseline parameter. We subsequently calculated the R^2 values of the predicted Gaussian fit for each voxel. The minimization procedure converged on a solution for all fits and initially no data were excluded following fitting (but see Results).

We additionally computed a suppression strength metric from the raw data, expressed in units of % signal change. In the whole-ROI analyses, we defined

suppression strength in each ROI as the difference between the averaged BOLD response in the orthogonal condition and the collinear condition. In the individual voxel analyses, suppression strength was defined as the difference between each voxel's average BOLD response in the orthogonal configuration and that in the collinear configuration.

Statistical tests were implemented in MATLAB and R using custom scripts. For statistical comparisons across observers, unless stated otherwise, we ran a repeated-measures ANOVA with visual area as grouping factor (observers were treated as random variables), and applied Bonferroni correction for the number of visual areas to all post-hoc t-tests. When conducting one-sample t-tests separately within individual ROIs, no corrections were applied. Unless stated otherwise, we report measures of dispersion as standard error of the mean (SEM).

Results

Tuned normalization bandwidth: whole-ROI analyses

Consistent with existing literature (Bloem & Ling, 2019; Chen, 2014; Flevaris & Murray, 2015; Joo et al., 2012; Ling et al., 2015; McDonald et al., 2009; Pihlaja et al., 2008; Poltoratski et al., 2017; Williams et al., 2003), collinear stimuli evoked lower amplitude BOLD responses compared to orthogonal stimuli. Suppression strength (the difference between averaged BOLD responses in the orthogonal condition and that in the collinear condition) was largely positive, except in a single ROI for one of the 10 observers. The average suppression strength was 0.23 ± 0.05 in V1, 0.28 ± 0.05 in V2, and 0.2 ± 0.05 in V3. There was a significant main effect of visual area ($F(2,18)$ = 6.559, p = 0.007); post-hoc t-tests revealed this reflected a significant difference between

suppression strength in V2 and V3 ($t(9) = 4.58$, $p = 0.004$). Suppression strength in each ROI differed significantly from 0 (one-sample t-tests, V1: $t(9) = 4.6$, $p = 0.001$, V2: $t(9) = 5.16$, $p < 0.001$, V3: $t(9) = 3.96$, $p = 0.003$). Importantly, for the majority of observers, and for all 3 visual areas V1/V2/V3, BOLD response magnitude showed a clear gradual increase as a function of orientation difference (**Figure 2.2**), indicating that suppression strength changes progressively as a function of the degree of orientation similarity. Note that as the data plotted in Figure 2 are normalized with respect to the response in the orthogonal condition, the values listed above do not directly correspond to the differences observed in Figure 2. To characterize the overall orientation-tuned normalization response, we averaged across all voxels in each ROI and observer, and fit the data individually (per observer and ROI) with a half-Gaussian function. An example of fits for a representative observer are shown in **Figure 2.3**. With these fits, we obtained an estimate of the standard deviation (bandwidth), amplitude and baseline for each observer and ROI. Note that due to the shape of the tuned normalization function, the fit is an inverted Gaussian, therefore decreasing in amplitude towards the mean. As a result, the baseline parameter represents the upper bound of the Gaussian fit, and the amplitude represents the difference between the baseline and the y-intercept of the curve (which corresponds to BOLD response magnitude of the collinear condition), and has a negative value for all observers. The amplitude parameter is closely related to suppression strength, as it is the difference between the fitted responses to collinear and orthogonal conditions, while suppression strength is the difference between the actual BOLD responses in those two conditions. Therefore, more negative amplitude estimates and

larger suppression strength point to greater suppression (larger difference between the collinear and orthogonal ends of the fitted Gaussian curves). The baseline is an additive parameter included to account for different absolute BOLD signal change magnitudes across observers and ROIs. We largely focus on our measures of suppression strength, and the Gaussian parameter estimates of bandwidth and amplitude for the remainder of the analyses.

The Gaussian fit the data well (average R^2 0.66 ± 0.11 in V1, 0.8 ± 0.08 in V2, and 0.73 ± 0.08 in V3), enabling us to rely on the standard deviation parameter as a proxy for tuned normalization tuning width estimates. Using this approach, we found that the average bandwidth was 42.3° ± 10.5° in V1, 32.7° ± 6.7° in V2, and 34.7° ± 7.6° in V3. Observer-averaged estimates of the parameters are displayed in **Figure 2.4A**. R^2 values are included in **Figure 2.5**. There were no between-ROI statistical differences in the bandwidth parameter ($F(2,18) = 1.422$, $p = 0.27$), or the amplitude parameter ($F(2,18) = 1.56$, $p = 0.24$). The baseline parameter was significantly different between visual areas ($F(2,18) = 13.18$, $p < 0.001$), reflecting lower BOLD responses in area V3 (1.07% ± 0.08% signal change) compared to V1 and V2 (1.36% ± 0.1%; $t(9) = 5.18$, $p = 0.002$ and 1.32% ± 0.1%; $t(9) = 3.82$, $p = 0.012$, respectively – see **Figure 2.3**). There was also considerable inter-observer variability in the bandwidth parameter. As can be seen in **Figure 2.4A**, the average is slightly inflated due to a few observers whose BOLD response did not show strong orientation-tuned suppression, resulting in a flat function.

Tuned normalization bandwidth: voxel-level analyses

Model fitting and voxel exclusion. We then carried out the fitting procedure at the individual voxel level for each observer and ROI. The median for each parameter estimate was computed and averaged across observers. Averaged median R^2 was 0.52 ± 0.07 in V1, 0.59 ± 0.07 in V2, and 0.52 ± 0.06 in V3, reflecting the increased noisiness in the single-voxel data, as compared to ROI averages. As with the ROI-averaged data, there were no significant differences between ROIs in the bandwidth parameter ($F(2,18) = 0.7$, $p = 0.51$). There was a main effect of ROI in the baseline parameter ($F(2,18) = 13.65$, $p < 0.001$), again reflecting differences between V1 (1.05% ± 0.07% signal change) and V3 (0.87% ± 0.07% signal change; $t(9) = 4.74$, $p = 0.003$) and V2 (1.07% ± 0.08% signal change) and V3 ($t(9) = 5.53$, $p = 0.001$). We also observed a main effect of ROI in the amplitude parameter ($F(2,18) = 5.57$, $p = 0.013$), which reflected marginally significant differences between V1 and V2 ($t(9) = 2.95$, $p = 0.049$) and differences between V2 and V3 ($t(9) = -3.38$, $p = 0.025$); amplitude was lower in V2 than in the other visual areas (-0.28 ± 0.04 in V2 vs. -0.23 ± 0.04 in V1 and -0.23 ± 0.04 in V3). There was large variation in the bandwidth estimates between observers, but also between individual voxels. Specifically, a number of voxels showed weak orientation-tuned suppression, meaning that their estimated sigma value fell exactly below the upper boundary of 90° and their BOLD response remained largely flat across orientation differences, compared to the remaining voxels. Upon further inspection, the voxels in question also showed overall lower R^2 values (V1: 0.26 ± 0.04, V2: 0.31 ± 0.06, V3: 0.28 ± 0.04), and weaker overall responses (lower BOLD % signal change; **Figure 2.6**). To obtain a clearer picture

of those voxels which did show tuning, we excluded a subset of voxels whose sigma estimate fell above 85°, capturing those that lacked tuning and had the maximum possible standard deviation (see **Figure 2.7** for the distribution of bandwidths across all voxels and observers). This led to the removal of 26.7% (± 4.5%) of voxels on average in V1, 22.7% (± 4.3%) in V2, and 26% (± 4.2%) in V3. We then computed within-ROI median and observer-average estimates from the remaining voxels, shown in **Figure 2.4B**. Voxel exclusion resulted in an overall increase in R^2 compared to a voxel-wise analysis with all voxels included (V1 average: 0.58 ± 0.06, V2: 0.64 ± 0.06, V3: 0.58 ± 0.06) and a decrease in the sigma parameter average, but did not qualitatively change the overall results, which found no differences between visual areas in bandwidth ($F(2,18) = 1.4$, $p = 0.27$). There was again a main effect of ROI in the baseline parameter ($F(2,18) = 12.4$, $p < 0.001$), with significant post-hoc t-tests in comparisons between V1 and V3 (V1: 1.05 ± 0.07, V3: 0.86 ± 0.08; $t(9) = 3.53$, $p = 0.019$), and V2 and V3 (V2: 1.09 ± 0.08; $t(9) = 4.74$, $p = 0.003$). The amplitude parameter also differed between areas ($F(2,18) = 4.7$, $p = 0.023$), this time with a significant post-hoc test only in V2 vs. V3 (V2: -0.26 ± 0.05, V3: -0.2 ± 0.04; $t(9) = -3.39$, $p = 0.024$). Since there was high similarity between the statistical comparisons in these two samples, subsequent analyses were carried out using the voxel subset with stricter selection criteria of SD < 85°. The tuning width (standard deviation) estimates obtained with this voxel subset were 23.1° ± 1.3° in V1, 24.4° ± 1.2° in V2, and 22.8° ± 1.5° in V3, calculated by averaging ROI-median estimates within ROI across observers.

Voxel-wise suppression strength was positive for the majority of voxels. The observer-averaged suppression strength (calculated as average of each observer's ROI medians) was somewhat lower compared to the whole-ROI averages; 0.2 ± 0.04 in V1, 0.23 ± 0.05 in V2, and 0.16 ± 0.04 in V3. There was again a significant main effect of visual area on suppression strength ($F(2,18) = 6.53$, $p = 0.007$), reflecting lower suppression strength in V3 compared to V2 ($t(9) = 4.27$, $p = 0.006$), results consistent with the ROI average. One-sample t-tests again showed that suppression strength in each ROI was significantly different from 0 (V1: $t(9) = 4.62$, $p = 0.001$, V2: $t(9) = 4.99$, $p < 0.001$, V3: $t(9) = 3.6$, $p = 0.006$).

As an additional analysis to validate the correspondence between amplitude and suppression strength, we conducted a Spearman correlation between voxel-wise amplitude estimates and suppression strength in each observer and ROI. The observer-averaged correlation coefficients were -0.85 (± 0.03) in V1, -0.87 (± 0.02) in V2, and -0.87 (± 0.02) in V3, indicating strong correspondence, where more negative amplitudes (representing larger differences between the fitted 90° response and the fitted 0° response) track larger suppression strengths.

Tuned suppression and voxel spatial preference. To further investigate properties of tuning width and suppression strength, we carried out additional exploratory analyses to characterize any dependence of suppression strength or the fitted parameters on voxels' retinotopic preference. To examine retinotopic preference dependencies, we binned the voxels into 8 eccentricity bins, logarithmically spaced between 1.5º and 8.5º (the inner and outer limits of the stimulus, respectively), collapsed across polar angle

preference (**Figure 2.8**). Following this, we computed the median estimate for each parameter per bin, observer and ROI. For each ROI and observer, we then fit a linear function through the data. Averages of the slope estimates were subsequently computed across observers. We assessed the relationship between the parameters and eccentricity by submitting the slopes for each ROI to a one-sample t-test. The slope for the sigma parameter was found to be significantly different from zero in all 3 ROIs (slopes: -1.2 ± 0.29 in V1, $t(9) = -4$, $p = 0.003$; -1.02 ± 0.31 in V2, $t(9) = -3.3$, $p = 0.009$; -1.14 ± 0.39 in V3, $t(9) = -2.9$, $p = 0.018$), indicating that the tuning width of tuned normalization decreases with increasing eccentricity. The averaged slopes for amplitude and suppression strength were not significantly different from zero in any of the ROIs.

We further explored whether these relationships also depend on polar angle, as previous work has shown that perceptual processing can differ between visual field segments, for example between upper and lower visual fields (Carrasco et al., 2001; Himmelberg et al., 2020; Kupers et al., 2019). We separated the visual field into quadrants, each comprising 90° above, below, left and right of fixation. Within each visual field segment, we averaged the median estimates across eccentricity bins for each observer and ROI to look for differences between visual field segments. As can be seen in **Figure 2.9**, there appeared to be significant differences in suppression strength between the upper and lower visual fields (paired t-tests; V1 $t(9) = 4.17$, $p = 0.002$; V2 $t(9) = 7.83$, $p < 0.001$; V3 $t(9) = 2.65$, $p = 0.026$), where suppression strength is generally stronger in the upper visual field (0.31 ± 0.06 in upper visual field vs 0.19 ± 0.04 in lower visual field); this relationship holds in all three ROIs. We also found amplitude

differences across the three ROIs between upper and lower quadrants: higher amplitude estimates were observed in the lower visual field across ROIs (paired t-tests; V1 $t(9) = -3.58$, $p = 0.006$; V2 $t(9) = -8.3$, $p < 0.001$; V3 $t(9) = -3.27$, $p = 0.01$). The average amplitude in the upper visual field was -0.35 ± 0.05 and -0.22 ± 0.05 in the lower visual field. As mentioned above, more negative amplitude estimates, together with larger suppression strength, indicate a stronger suppression effect in this area of the visual field. Generally, we did not observe reliable differences in bandwidths between visual field segments; only in V2, in the comparison between left and right visual field, was there a statistically significant difference ($t(9) = -3$, $p = 0.015$), with wider sigma in the right visual field. Plots depicting slopes between the left and right quadrants of the visual field are shown in **Figure 2.10**). Additionally, we have included surface maps for polar angle, eccentricity and bandwidth for our representative observer in **Figure 2.11**.

In an additional exploratory analysis, we asked whether the position of a voxel on the wedge had an influence on suppression strength. Most differential suppression in our stimulus was thought to occur at the boundary regions between adjacent wedge components; each wedge serves as a "center" as well as a "surround" in this circular full-field arrangement. One could therefore assume that suppression strength might differ between portions of our participants' visual fields corresponding to the boundaries between wedges and the inner surfaces of the wedge components – specifically, that we would find tuning by orientation difference at the boundaries of the wedges but not on the surfaces of the components where the orientation content is always collinear. To explore this possibility, we binned voxel polar angle estimates from the pRF analysis according to

their position on the wedge, into 1° bins, as follows: first, we split each wedge into 18 bins (as each wedge spanned 18°) and grouped voxels from corresponding bins in each of the 20 wedges together. We then grouped them based on the distance from the boundary, as opposed to just position on the wedge; for example, voxels from the 0-1-degree bin and 17-18-degree bin were now in the same group, as both are equally distant from the boundary of the wedge. This step produced 9 bins (0-1-degree from boundary, up until 8-9-degree from boundary, which was the center of the wedge). For each ROI and observer, we then took the median of estimates for all voxels in each bin, and averaged over observers for each ROI. We then fit a simple linear model to the observer-averaged data in each ROI. Distance from boundary did not act as a significant predictor for suppression strength or any of the Gaussian parameters (See **Figure 2.12**). That is not to say there is no potential spatial gradient, but our particular wedge and border sizes were likely too small given the estimated pRF sizes of our voxels, to pick up on these presumably finer scale gradients in suppression strength. We designed our stimulus to elicit the most suppression on a large population scale, and to maximize the number of voxels differentially modulated by suppression throughout the display. Thus, the wedges took up 0.47 dva at the innermost stimulus eccentricity and 2.66 dva at the outermost eccentricity (see Methods), while the "gaps" between the components took up less than 0.1 dva at all eccentricities. On the other hand, the observer-averaged median estimates of voxel receptive field sizes were much larger than the gap sizes (in the closest eccentricity bin – V1: 0.43 ± 0.01 dva, V2: 0.53 ± 0.05 dva, V3: 0.61 ± 0.03 dva; in the highest eccentricity bin – V1: 0.75 ± 0.1 dva, V2: 1.6 ± 0.06 dva, V3: 2.1 ± 0.13 dva).

A final exploratory analysis examined potential causes of broad bandwidth estimates in the excluded voxels, namely that they may be caused by noisy voxels, as indicated by the Gaussian R^2. As expected, there were significant negative correlations between bandwidth and R^2 (Spearman correlation, observer-averaged correlation coefficients -0.25 ± 0.03 in V1, -0.23 ± 0.04 in V2, and -0.24 ± 0.03 in V3); this agrees with our above observation of overall lower average R^2 values in broad-bandwidth voxels. We first looked at the spatial distribution of these voxels across the stimulus display. Broad bandwidth voxels were found across the entire visual field, as opposed to being bound to e.g. only high eccentricities, suggesting that broad bandwidth fitting in these voxels was not an artifact of their visual field position. We subsequently examined the behavior of bandwidth and R^2, in all the voxels that underwent the Gaussian fitting, across the visual field; if broad bandwidths were simply a product of lower R^2, we would expect corresponding behavior of the two parameters – i.e., as bandwidth becomes sharper towards the periphery, we would expect higher R^2 values there. However, this was not the case; we binned R^2 and bandwidth for the whole voxel set into eccentricity bins, following the same steps we took in creating Figure 5. While R^2 tended to decrease with eccentricity (slope was significantly different from 0 in V3, $t(9) = -2.62$, $p = 0.028$), bandwidth showed a similar pattern as in Figure 5, with slope significantly different from 0 in V1 ($t(9) = -3.24$, $p = 0.01$), and no increases in slope. This suggests that broad sigma estimates are not entirely a product of low R^2.

Discussion

Our results suggest that the bandwidth of tuned suppression in human early visual cortex is around 23º on average, and between 20º and 30º for most observers. This finding, which to our knowledge is the first fMRI investigation of tuned suppression bandwidth, expands on previous studies showing reliable differences in BOLD signal amplitude between collinear and orthogonal stimulus configurations (Bloem & Ling, 2019; Chen, 2014; Flevaris & Murray, 2015; Joo et al., 2012; Ling et al., 2015; McDonald et al., 2009; Pihlaja et al., 2008; Poltoratski et al., 2017; Williams et al., 2003).

These findings build on existing animal and psychophysical work, some of which did explore the extent of orientation-dependent tuning of surround suppression. One animal study (Trott & Born, 2015) varied orientation differences in a center-surround stimulus, gradually changing the relative orientations of the two components in 30º increments, and found that response magnitudes decrease incrementally with increasing similarity – although suppression bandwidth was not directly investigated. Shushruth and colleagues (Shushruth et al., 2013) investigated suppression tuning more directly and discovered significant tuning with center-surround stimuli, which they roughly matched between macaque V1 recordings and an accompanying human psychophysical study using contrast matching. Tuning was sharper when the center and surround were in proximity, and weaker when they were far apart. Petrov and colleagues (Petrov et al., 2005) also used a contrast matching task with center-surround stimuli in humans and reported that surround suppression was quite tightly tuned, with strong suppression at similar orientations and almost no suppression at orientation differences that exceed 45º.

Cannon and Fullenkamp (Cannon & Fullenkamp, 1991) concluded there are two tuning mechanisms in their study using a psychophysical contrast matching task with human observers and center-surround stimuli: one suppression mechanism was narrowly tuned, disappearing at an orientation difference of about 15º, while the other decreased more gradually and was still present when the center-surround orientation difference reached 90º. Our stimuli are not center-surround, and so we were unable to employ a center-only condition to compare the amount of suppression at 90º with a no-surround configuration. At the same time, our minimum orientation difference was 15º (although see discussion on stimulus limitations below), and so we cannot study the narrower tuning mechanism. However, the BOLD signal strength in our dataset continued to change with increasing orientation differences up until a difference of 90º, suggesting some similarity between the mechanism we captured and the broader tuning component reported by Cannon and Fullenkamp (Cannon & Fullenkamp, 1991). More recently a psychophysical study (Phillips et al., 2021) examined the bandwidth of suppression in motion direction tuning by using center and surround grating stimuli which differed in relative orientation as well as motion direction, and found that when motion directions of the center and surround matched, suppression strength was at its maximum, decreasing with increasing angular direction differences. The reported average tuning width for motion direction differences was 28º, which bears close resemblance to our average estimates of tuning for orientation differences.

When comparing our suppression strength results to existing literature, it is important to note that the work presented here measures only the tuned component of

suppression, i.e., the difference in BOLD response between collinear and orthogonal stimulus configurations, whereas most literature examining suppression strength focuses on contrasting BOLD responses to a central stimulus with or without a surround (i.e., untuned suppression). Untuned suppression was generally found to be stronger in extrastriate areas, as compared to V1 (Kastner, 1998; Zenger-Landolt & Heeger, 2003). Few groups directly examined differences in tuned suppression between visual areas. Bloem and Ling (Bloem & Ling, 2019) found a decrease in tuned normalization magnitude across early visual areas, while Poltoratski et al. (Poltoratski et al., 2017), who used flanker displays, reported no statistical differences. Other studies examining tuned normalization did not conduct direct statistical comparisons between areas (Joo et al., 2012; McDonald et al., 2009), although one reported that tuned suppression was significantly different from zero in extrastriate cortex but not in V1 (McDonald et al., 2009).

Likewise, in previous psychophysical work, tuned and untuned suppression have shown diverging eccentricity effects. Untuned suppression strength typically increases with eccentricity with a plateau at around 4 dva from fixation (Petrov et al., 2005; Xing & Heeger, 2000). However, Xing and Heeger (Xing & Heeger, 2000) found no eccentricity-based differences in the magnitude of tuned suppression, but noted that suppression was stronger, and less specific, in the visual field periphery. Our neuroimaging results also revealed no eccentricity effects on tuned suppression strength. Unlike previous studies, we did not use a cortical magnification factor to adjust the spatial frequency of our stimuli with eccentricity, which could be a contributing factor for some of the differences

between previous findings and our results. The other obvious difference is the overall stimulus configuration – all previous work described above used center-surround grating stimuli or flanker displays to elicit surround suppression. While our stimulus clearly elicits measurable orientation-tuned responses in the visual cortex, it makes direct comparisons with studies using different stimuli difficult.

A limitation of our stimulus was the orientation filter bandwidth (10°, see Methods). The addition of a filter bandwidth means that the possible orientation differences between individual components could range *around* the intended value (for instance, in a 15° orientation difference condition where one component is 0° and the other 15°, the filter bandwidth used means that the actual minimum difference between some of the individual components could range between 5° and 25°). While this clearly does not hinder our ability to measure orientation-tuned suppression, it imposes a limit on the precision of bandwidth estimates.

The goal of the present study was to estimate the bandwidth of orientation-tuned suppression, which is distinct from orientation tuning bandwidth. That said, our ability to measure orientation-tuned suppression bandwidth is also limited by the underlying *orientation tuning* in the visual areas of interest. Orientation tuning bandwidth within human visual cortex has been studied with psychophysics and fMRI. For instance, one study (Tootell et al., 1998) used fMRI adaptation to measure transient BOLD responses to orientation changes in a full-field grating stimulus, and estimated the bandwidth of orientation tuning in human V1 at 45°. In psychophysical work using the orientation noise-masking technique (Ling & Blake, 2009), estimated orientation tuning bandwidth

ranged between 15 and 30 degrees. Thus, similar to investigations of tuned normalization bandwidth, estimates of orientation tuning bandwidth in human early visual cortex vary quite widely.

Our results could be of use in further constraining current models of divisive normalization so that they take more granular feature similarity into account. The notion that suppression strength varies with feature similarity has been explored in research investigating how the visual cortex appears optimized for natural scene properties. According to theories of efficient coding, the visual system attempts to remove redundancies in incoming visual input to achieve a more efficient representation, and one of the ways this could be accomplished is through divisive normalization (Schwartz & Simoncelli, 2001). Indeed, there is evidence that surround suppression among V1 neurons in the macaque is stronger for homogeneous, as opposed to heterogeneous, images, suggesting that feature tuning supports efficient coding of visual stimuli, discounting less informative homogeneous regions and highlighting contrasts between objects to promote figure-ground segregation (Coen-Cagli et al., 2015; Vinje & Gallant, 2000). Coen-Cagli and colleagues (Coen-Cagli et al., 2015) further demonstrated the significance of feature similarity in divisive normalization by showing that the standard normalization model only accounts for roughly half the variance in macaque V1 neural responses to natural images, but when a gating component based upon the degree of homogeneity in the images was added to the model, the variance explained dramatically improved. In this flexible model variant, when the center and the surrounding portions of an image share their features and are thus considered homogeneous, the surround suppression

mechanism is engaged; however, when the two components are judged as heterogeneous, the surround influence is "switched off" and the neural response to the center component of the visual stimulus is not normalized by the response to the surrounding area. While this model improves data fits significantly, incorporation of our more fully characterized surround suppression function could provide an even closer prediction; having knowledge of the bandwidth could allow for a gradual adjustment of suppression strength, and yield an estimate of how much suppression one might expect for a particular orientation difference.

More evidence to support the idea that tuned normalization strength reflects the most common properties of natural scenes comes from natural image statistics analyses. While it was well established that nearby contours are most likely to be collinear (Geisler et al., 2001), Sigman and colleagues (Sigman et al., 2001) examined the number of occurrences of orientation differences graded in 11.25º steps between image segments, found in a database of 4,000 photographs of natural scenes. Pairs of segments containing orientations collinear to each other occurred most frequently, and there was a gradual drop in frequency with decreasing orientation similarity, with orthogonal orientation segment pairs being least likely to co-occur in a scene, matching our results. This finding squares with our results particularly in light of efficient coding perspectives. The surround region can be thought of as setting the context for the center (Joo et al., 2012; Schwartz & Simoncelli, 2001), and when the center matches this expectation (i.e., is collinear), the output of the units responding to the center is suppressed to reduce redundancy in neural coding. Conversely, when the center is orthogonal to the surround,

reduced suppression will ensure that this informative signal is processed with increased saliency. This supports the notion that suppression strength dependence on orientation similarity is well matched to the statistics of natural stimuli; more predictable, redundant components, which are also more likely to co-occur in scenes, are suppressed, while the less frequently occurring stimuli, which often highlight regions of interest such as figure-ground transitions, benefit from increased saliency brought about by reduced suppression.

In summary, our results reveal that suppression strength in human early visual cortex depends on orientation differences between scene elements. Of course, other features play a role in our visual system's ability to parse visual scenes, such as similarities in contrast and spatial frequency, or distance between scene components (Bell & Badcock, 2008; Cannon & Fullenkamp, 1991; Coen-Cagli et al., 2015; Maloney & Clifford, 2015). Here we focused solely on orientation differences, but going forward, incorporating other dimensions, in particular stimulus contrast, would be a crucial next step in fully characterizing the profile of feature-tuned normalization.

Figures

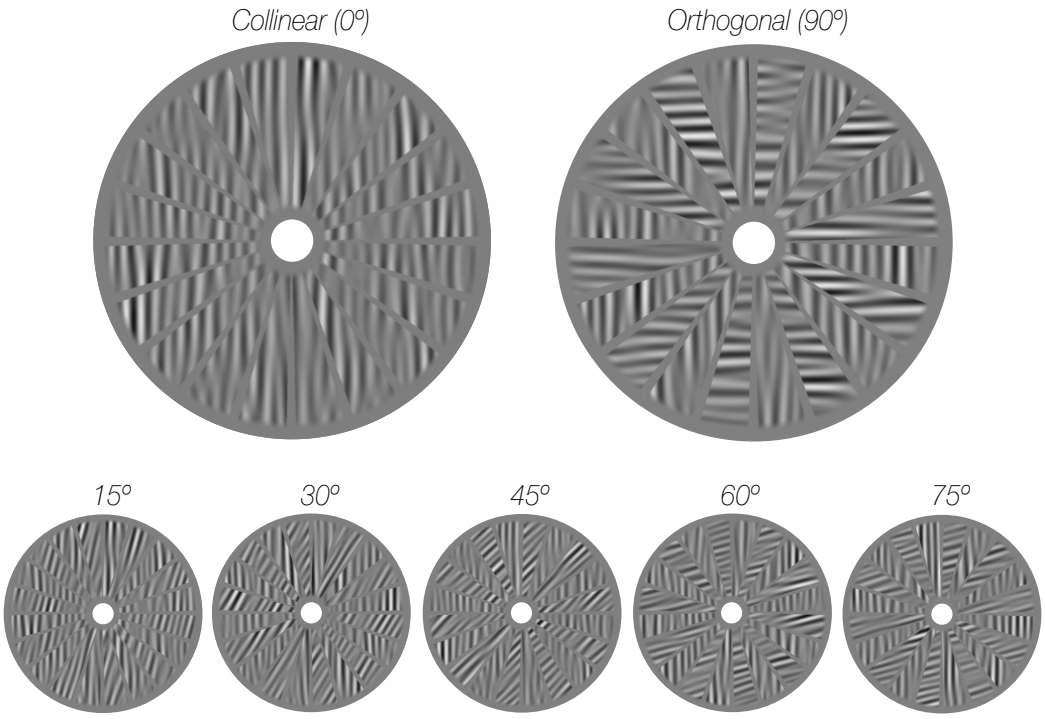

Figure 2.1 Experimental stimuli.

Examples of bandpass filtered stimuli used in the study. The full-field stimuli were composed of two interleaved sets of 10 wedges, filled with two independent noise patches. Stimuli are modified (increased contrast and decreased spatial frequency) for illustration purposes.

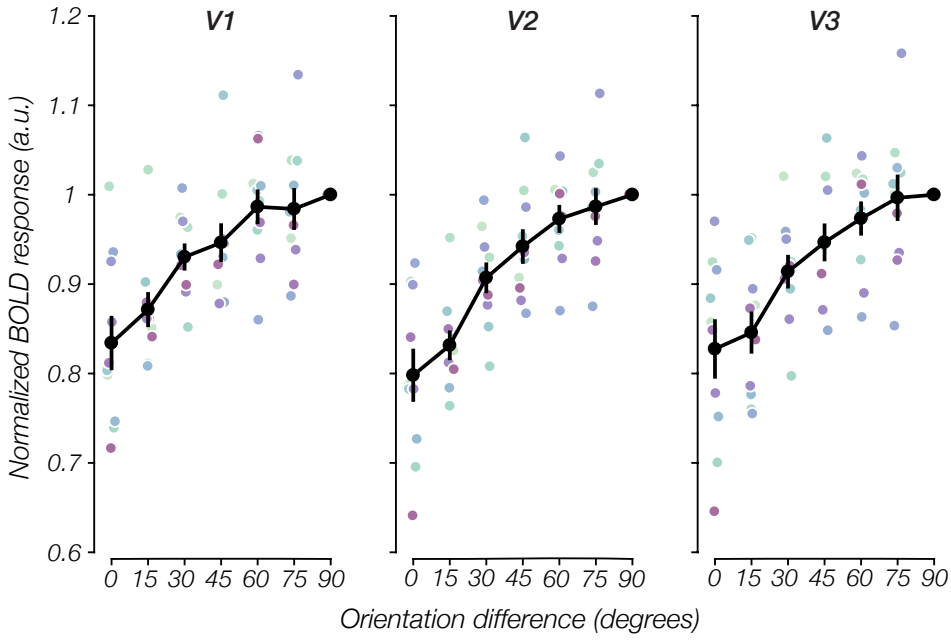

Figure 2.2 Observer-averaged BOLD responses by orientation difference.

BOLD responses as a function of orientation difference, across observers (n = 10), from V1-V3. Responses for each observer were normalized with respect to the response to the 90º orientation difference condition. Thick black line represents the observer average of these normalized responses. Colored points depict responses for individual observers. Error bars represent ± 1 SEM.

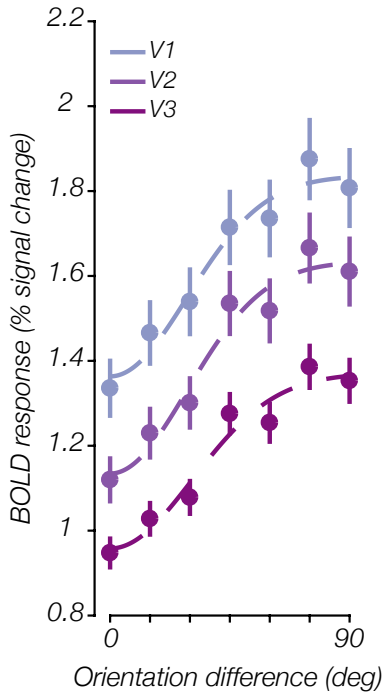

Figure 2.3 Example observer data and Gaussian fits.

Gaussian function fit to the tuned normalization data for a single example observer. Points indicate the average BOLD signal for each ROI ± 1 SEM. Dashed lines represent the Gaussian fits for each ROI for this representative participant.

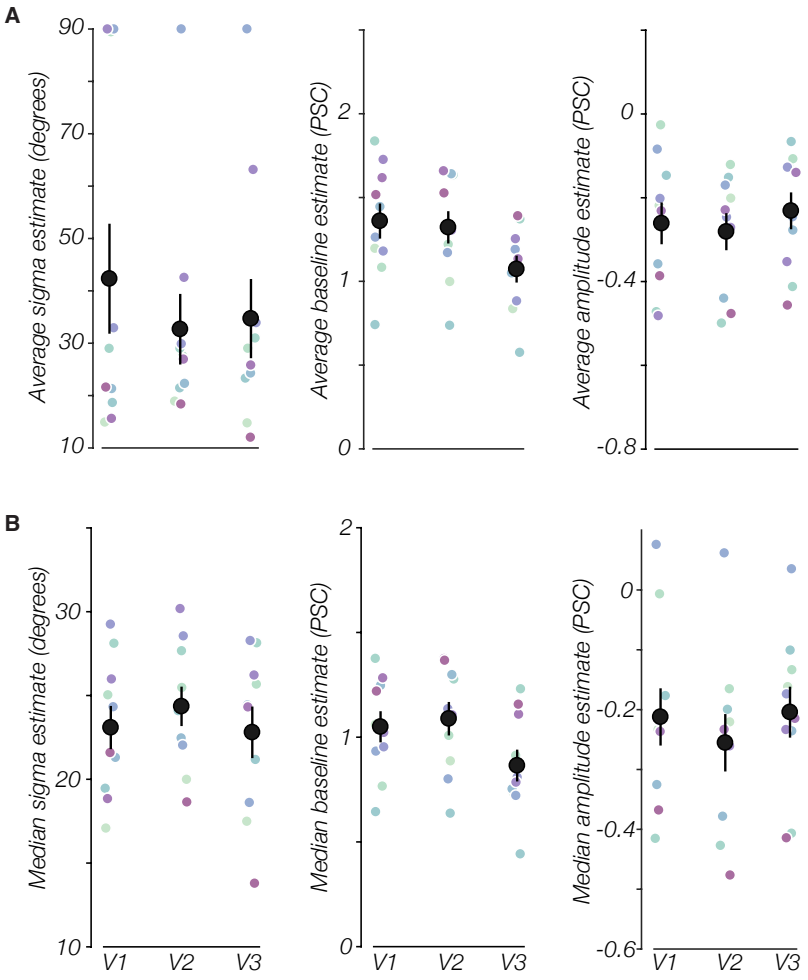

Figure 2.4 Gaussian fit parameter estimates across visual areas.

(A) Parameter estimates (n = 10) for all visual areas, obtained by fitting ROI-averaged voxel activation values for each observer and ROI (PSC = % signal change). (B) Median (per observer/ROI) parameter estimates for all visual areas following removal of voxels whose sigma parameter was estimated to be wider than 85º (n = 10). Data are shown as observer averages ± 1 SEM. Colored points represent individual observer parameter estimates while large black circles represent the observer averages.

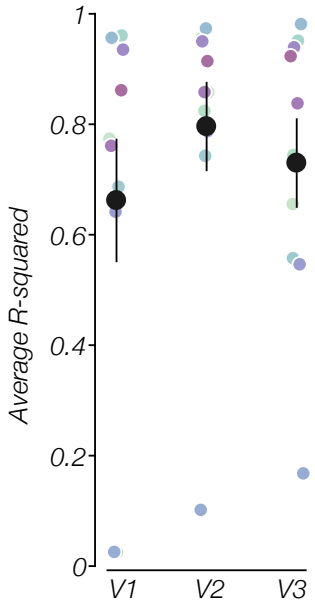

Figure 2.5 Observer-averaged R^2 values of the Gaussian fit functions.

Observer-averaged (n = 10) R^2 values of the Gaussian fit functions, displayed for the whole-ROI fitting. Error bars represent ± 1 SEM. Black points represent between-observer averages, colored points represent individual observer R^2 values.

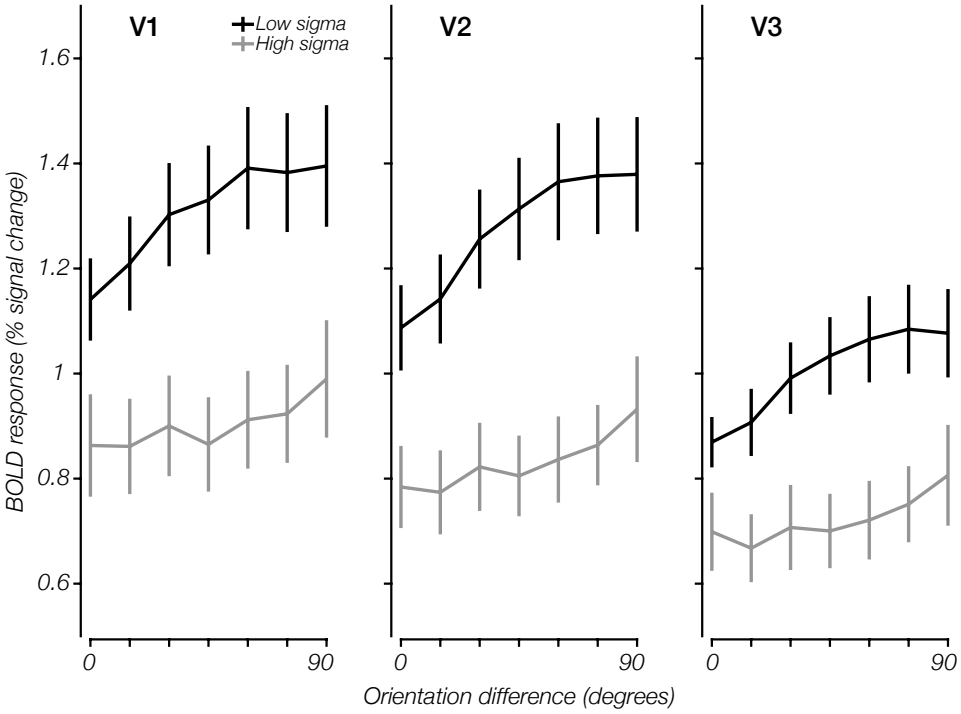

Figure 2.6 Averaged BOLD responses for voxel subsets based on bandwidth.

Averaged raw data for the subsets of voxels including and excluding voxels with sigma estimates above 85°. Data are shown as observer-averaged (n = 10) BOLD signal change in each orientation difference condition ± 1 SEM.

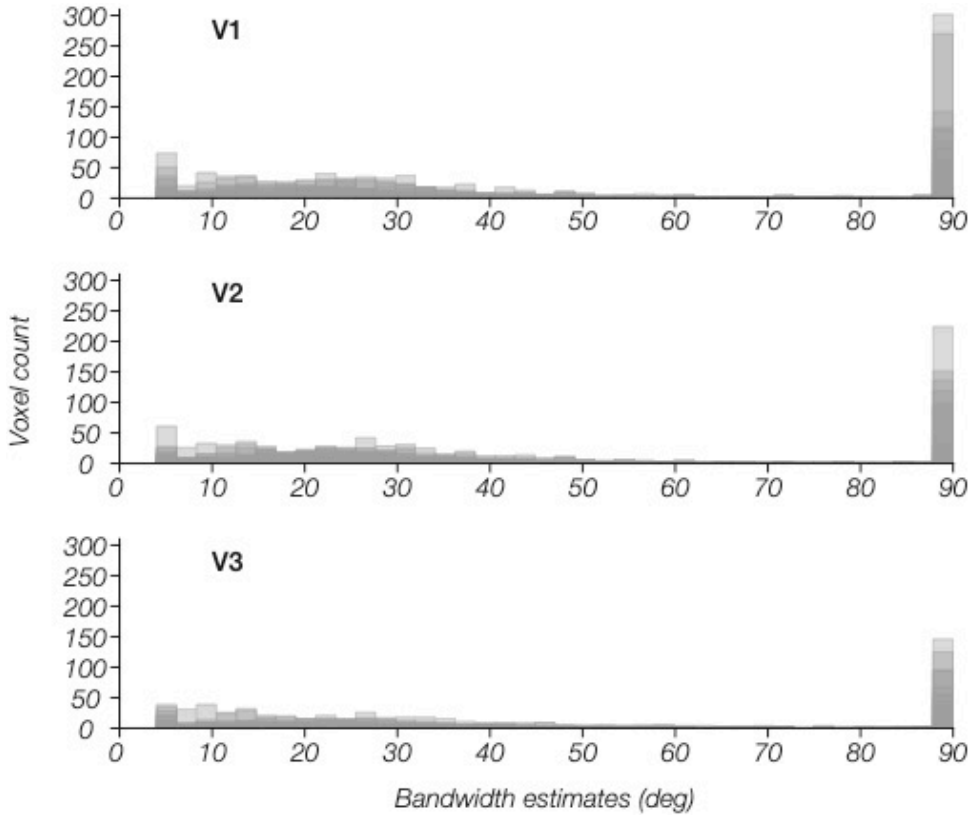

Figure 2.7 Gaussian bandwidth estimate distribution across visual cortex.

Histograms showing distributions of bandwidth estimates for all voxels that underwent Gaussian fitting in all observers (n = 10). The histograms for all observers were superimposed for each ROI.

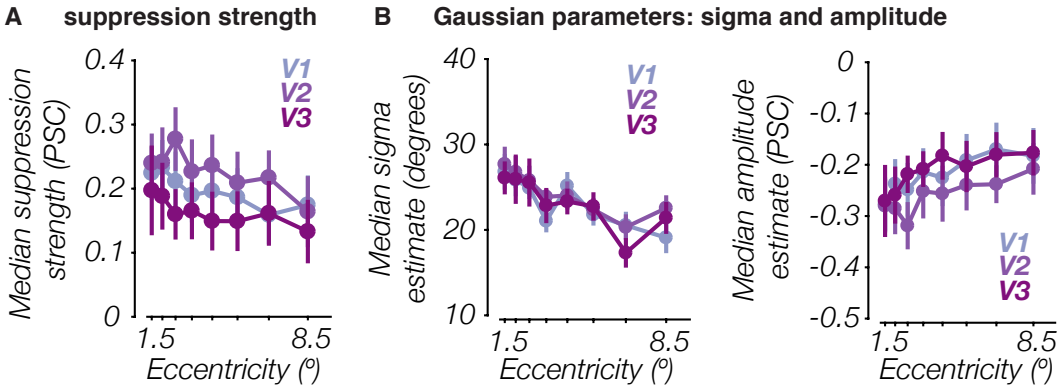

Figure 2.8 Relationship between tuned normalization and eccentricity.

Relationship between (A) data-derived suppression strength and eccentricity (n = 10, PSC = % signal change) (B) the Gaussian parameter estimates sigma (middle) and amplitude (right) vs. eccentricity (n = 10). The data points represent observer mean of the parameters ± 1 SEM, per ROI and eccentricity bin.

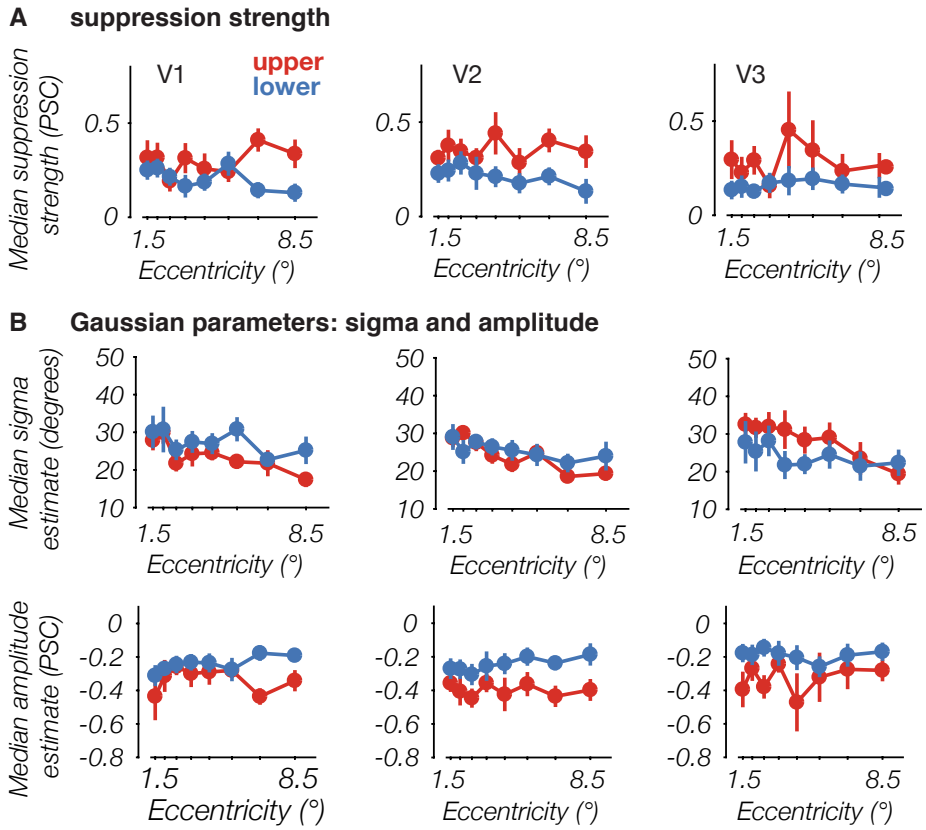

Figure 2.9 Tuned normalization differences between upper vs. lower visual field.

(A) median suppression strength (n = 10, PSC = percent signal change), (B) standard deviation (top) and amplitude (bottom), n = 10. Points represent mean across observers ± 1 SEM plotted as a function of eccentricity, separated by visual field segment. Each panel shows the comparison between upper visual field (upper; red) and lower visual field (lower; blue).

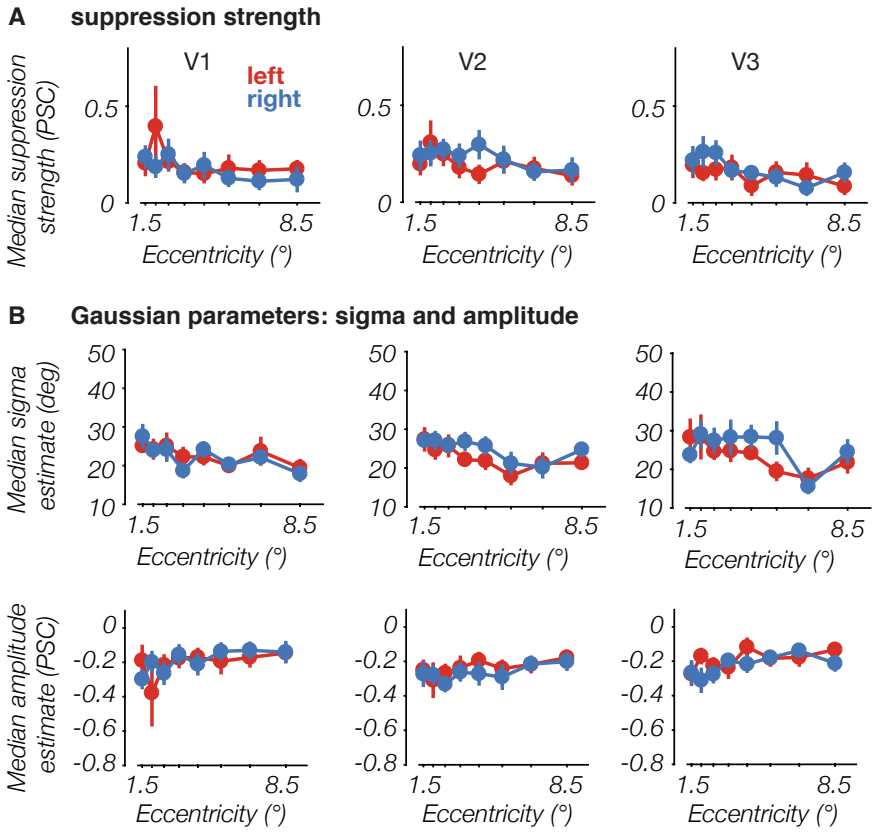

Figure 2.10 Tuned normalization differences between left vs. right visual field.

(A) median suppression strength (n = 10), (B) standard deviation (middle) and amplitude (bottom). Points represent mean across observers ± SEM (n = 10) plotted as a function of eccentricity, separated by visual field segment. Each panel shows the comparison between left visual field (left; red) and right visual field (right; blue). PSC = percent signal change.

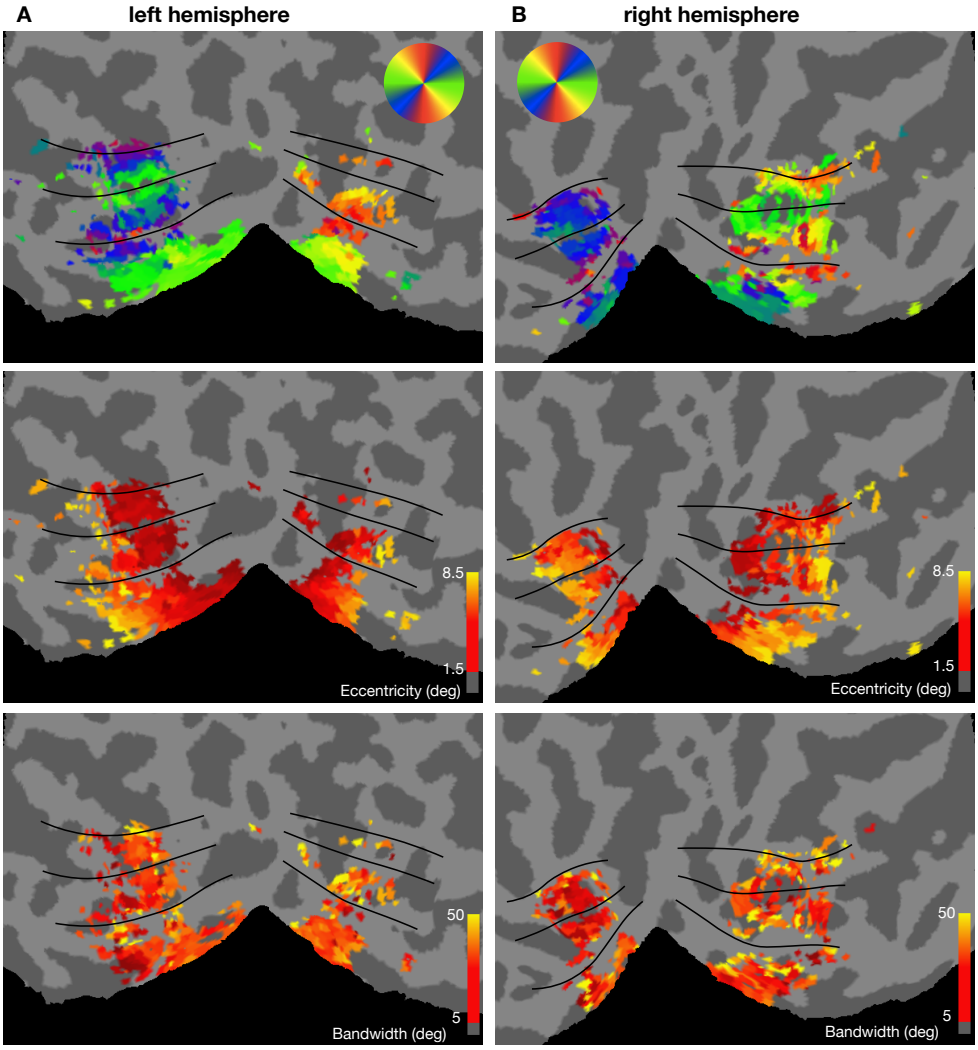

Figure 2.11 Example observer surface maps with pRF indices and tuned normalization bandwidth.

(A) left and (B) right hemisphere occipital cortex surface flat maps for the representative observer whose whole-ROI fit was shown in Figure 3 of the main manuscript. In the top panel, color indicates polar angle – inset color wheel represents the visual field and indicates the location of horizontal and vertical meridians. In the middle panel, color indicates eccentricity, and in the bottom panel, color indicates bandwidth (voxels whose bandwidth fell between 50 and 85 degrees are depicted in yellow). The black lines represent borders between early visual areas: bottom = V1, middle = V2, top = V3.

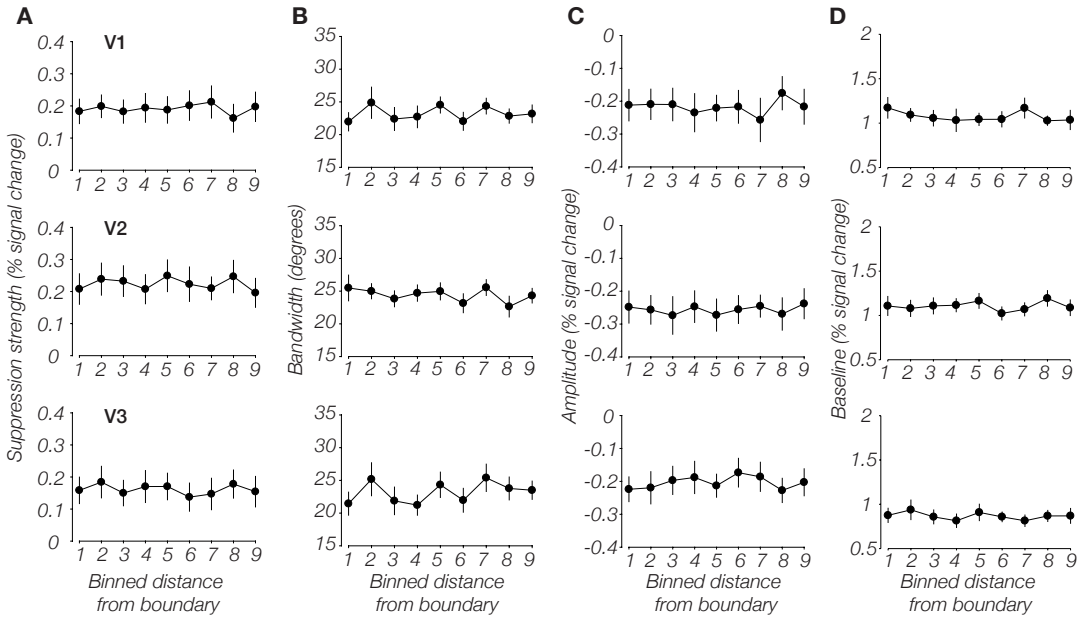

Figure 2.12 Suppression strength and Gaussian parameters as a function of voxel position in stimulus display.

(A) Suppression strength and (B-D) Gaussian parameters (B: bandwidth, C: amplitude, D: baseline) as a function of binned distance from wedge boundary. Data are shown as average across observers (n = 10) ± 1 SEM. Wedges spanned 18°. Bin 1 represents voxels found between 0°-1° away from boundary and bin 9 represents voxels found between 8°-9° away from boundary (i.e., center of the wedge).

CHAPTER THREE: ATTENTION PRESERVES THE SELECTIVITY OF FEATURE-TUNED NORMALIZATION

Introduction

Although our senses are constantly flooded with stimulation, they are surprisingly effective at filtering incoming information. To do so, the visual system deploys a number of strategies, some of which involve top-down, endogenous guidance, and others that are bottom-up and largely automatic. One such bottom-up process is *divisive normalization*, wherein the magnitude of a neuron's response to a stimulus is divided by the pooled responses of itself and neighboring neurons. Widely believed to be a canonical neural computation, normalization operates throughout the visual processing hierarchy (Carandini & Heeger, 2012), and offers a computational account of commonly observed phenomena such as the saturation of responses at high stimulus intensities, as well as surround suppression (Carandini & Heeger, 2012; Cavanaugh et al., 2002a).

Surround suppression, one of the key signatures of normalization, has been observed with electrophysiology (Cavanaugh et al., 2002a; DeAngelis et al., 1992; Self et al., 2014; Webb et al., 2005), as well as with large-scale population measures such as fMRI (Flevaris & Murray, 2015; Joo et al., 2012; Zenger-Landolt & Heeger, 2003) and EEG (Schallmo et al., 2019). Interestingly, not all surrounding stimuli and configurations are created equal: normalization appears to be feature-tuned, wherein the magnitude of suppression depends on the degree of feature similarity between the center and surrounding stimuli. Specifically, surrounding stimuli whose features (e.g. orientation) differ from the central stimulus are less suppressive than surrounding stimuli that share

identical properties with the center (Bloem et al., 2018; Bloem & Ling, 2019; Chen, 2014; Ling et al., 2015; McDonald et al., 2009; Petrov et al., 2005; Pihlaja et al., 2008; Poltoratski et al., 2017; Schallmo & Murray, 2016; Self et al., 2014; Shushruth et al., 2013; Trott & Born, 2015; Webb et al., 2005; Williams et al., 2003; Xing & Heeger, 2000). This feature-tuned component of normalization is proposed to play a key functional role in efficient coding during natural scene perception: by compressing representations of redundant information and enhancing the salience of unique contours, a figure can be better segregated from its background (Coen-Cagli et al., 2015; Schwartz & Simoncelli, 2001). Interestingly, the selectivity of feature-tuned suppression within human visual cortex appears to be nicely tuned to the image statistics of natural environments: by characterizing the bandwidth of orientation-tuned suppression in V1-V3, we found that bandwidths fell between 20-30 degrees orientation difference (Klímová et al., 2021), closely matching the distribution of orientation co-occurrences in natural scenes (Sigman et al., 2001).

Top-down processing also plays a key role in the moment-to-moment filtering of information. Visual attention helps parse scenes and select relevant stimuli for preferential processing, at the expense of unattended stimuli. This essential top-down mechanism allows us to regulate between the overabundance of sensory stimulation in our environment, and the limited processing resources available to the brain at any given moment (Desimone & Duncan, 1995). Directing attention to a stimulus is well known to increase the perceptual salience of its representation (Carrasco et al., 2004; Carrasco & Barbot, 2019; Yeshurun & Rashal, 2010) – a behavioral consequence of attention-driven

changes in the gain of neural responses (Kastner, 1998; Somers et al., 1999; Treue, 2001). While divisive normalization and attention independently regulate incoming visual information, evidence gathered in recent years suggests that they may be closely intertwined. Normalization models of attention propose that the gain changes elicited by attention result in altering the balance between excitation and inhibition, inherent in normalization, and effectively providing a release from suppression (Lee & Maunsell, 2009; Ni et al., 2012; Ni & Maunsell, 2017; Reynolds & Heeger, 2009). Functional neuroimaging in humans has found a similar link between tuned normalization and attention, wherein subpopulations exhibiting stronger tuned normalization also possessed stronger attentional modulation (Bloem & Ling, 2019). Consistent with this, psychophysical work has also found weaker attentional effects for stimuli that evoked weaker surround suppression (orthogonal arrangement as opposed to collinear; (Kınıklıoğlu & Boyaci, 2022).

Given this tight relationship between attention and normalization, could attention, in turn, also enhance the specificity of tuned normalization? To test this, we used fMRI to examine whether visual attention qualitatively modulates the profile of orientation-tuned normalization. While the normalization model of attention proposes that attentional benefits emerge from a release from normalization, it remains unclear whether this applies to both the tuned and untuned components of normalization. On the one hand, if attention tightened the bandwidth of orientation-tuned normalization, one might experience heightened sensitivity to smaller changes in orientation, which could enhance scene processing. Such a modulation would qualitatively change the way feature

differences in a scene are processed and perceived. On the other hand, work from our group and others (Klímová et al., 2021; Phillips et al., 2021; Sigman et al., 2001) has supported the notion that normalization bandwidths are tuned to the statistical regularities of our visual environments – larger differences in orientation and other visual features are highlighted (by inducing larger neural responses), while continuities in scenes and objects are discounted (Coen-Cagli et al., 2015). Given that tuned normalization already serves to faithfully adapt visual responses to the predominant distribution of orientations in the environment, perhaps attending to a stimulus should not augment this tuned function. To test these two possibilities in the human visual cortex, we leveraged full-field stimuli previously designed to measure the bandwidth of orientation-tuned suppression (Klímová et al., 2021) under two conditions: attending to the oriented stimuli and passive viewing (attending to a task at fixation). Our results revealed an overall enhancement of the BOLD response when the stimulus was attended to, but no discernible qualitative attentional effects on the orientation-difference tuning bandwidth, suggesting that while attention increases the neural responses to a stimulus, it does not alter properties of visual processing, which are likely titrated to natural scene statistics.

Materials and Methods

Observers

10 observers (4 female) took part in the study. All received monetary compensation for their time, except three who were also the authors of the study. The imaging for the attentional task was conducted in a single 2-hour session. An additional fMRI session was conducted prior to the study in order to obtain a high-resolution T1-

weighted scan and to conduct population receptive field (pRF) mapping to identify early visual areas V1-V3. All observers reported their vision as normal or corrected-to-normal and were between the ages of 18 and 40. All provided written informed consent and the study was approved by the Boston University Institutional Review Board.

Apparatus and Stimuli

Visual stimuli were generated with Psychophysics Toolbox in MATLAB (R2015b) on a MacBook Pro (OS X 10.7) and displayed in the scanner on a rear-projection screen using a linearized projection system (VPixx Technologies PROPixx DLP LED projector). The stimulus **(Figure 3.1B)**, subtending 3-17 degrees of visual angle (dva) in diameter was circular and composed of 20 wedges (two interleaved sets of 10 with variable orientation content). The spatial pattern filling each wedge was created by band-pass filtering white noise. The filter had an orientation bandwidth of 10° (5° on either side of the target orientation), and a spatial frequency bandwidth of 2-3 cycles per degree (cpd). Neighboring wedges could contain one out of six orientation differences, spanning between 0° (collinear) and 90° (orthogonal), in steps of 18°. Furthermore, each orientation difference could be offset by one of five "base orientations", which spanned between 9° and 81° (also in steps of 18°). Stimuli were always filtered with a vertical orientation, and on each trial, they were rotated by the value of the base orientation for that trial, which was counterbalanced across the scanning session (see below). All stimuli had 50% Michelson contrast and were presented on a mean luminance background. Each fMRI run contained 36 stimulus presentations, and each condition (combination of orientation difference and attentional state) was presented 3 times within a run.

Base orientations were fully counterbalanced across every 10 runs, such that each base orientation was presented 6 times for each orientation and attention condition across the 10 runs. Furthermore, we counterbalanced which set of wedges contained the base orientation and which would contain the offset orientation (3 out of 6 presentations each).

Experimental Procedure

Each observer completed 9-11 scan runs of the main task. The main task (408 TRs, 1 s TR) was a slow event-related design with a 6 s stimulus duration and variable ITI (between 4-14 s). Event scheduling was done using FreeSurfer's Optseq2 scheduling tool (Dale, 1999). The event schedule in each run was identical for all observers in terms of orientation difference and attention condition, but the order of base orientations on each trial differed between observers. Each main task run began with an 8-second baseline period, only showing the mid-luminance screen and a rapid letter stream within the fixation circle (see below). Observers were asked to maintain central fixation throughout the experiment.

A small fixation point was presented in the center of a white fixation circle (3 dva diameter). In front of the fixation point, white letters (letter size 0.75 dva) were presented in a rapid visual stream at 5 Hz (RSVP task). The fixation circle was always present, as was the letter stream; both remained white throughout the scan. The color of the fixation point differed between baseline, the attend-to-stimulus condition, and the attend-away condition, serving as a cue to the observer to indicate which task to perform. Throughout baseline/null periods, the fixation point was red. 500 ms before each 6-s task trial, the fixation point changed color to indicate the task for the upcoming trial. There was also a

500 ms letter cue, interrupting the RSVP stream, prior to each trial. A green fixation point and the cue letter “C” indicated that in the upcoming 6-second trial the observer will be asked to detect transparent color patches within the stimulus display and ignore the RSVP letter stream, so that covert attention is directed to the oriented stimulus. A black fixation point and the cue letter “F” told the observer that they should attend to the letter stream at fixation and monitor for target letters, thus passively viewing the stimulus while directing their attention away from it. The fixation point kept its task-specific color throughout the 6-second task block, to minimize the possibility that observers might erroneously focus on the wrong task.

During the attended trials **(Figure 3.1A, top four time points)**, observers were to maintain central fixation but monitor the entire stimulus for the appearance of a faint color patch that could appear over any of the wedges. The color patch appeared gradually, ramping up from alpha (opacity) = 0 to each observer’s individual alpha level (see below) and then back to 0 in the span of one second; the ramp-up took 400 ms, maximum alpha remained for 200 ms, followed by 400 ms ramp-down. A color target appeared with 60% probability after two consecutive non-target stimulus refresh cycles (100 ms). This led to between 0-2 color targets appearing in each 6-s task block. The participants held a two-button box throughout the experiment with instructions to press the left key if they saw a red patch, and the right key if they saw a blue patch.

In the passive viewing trials **(Figure 3.1A, bottom three time points)**, participants monitored the central fixation letter stream, in which a new letter was presented every 200 ms. Upon seeing the letter “J”, participants were to press the left

button, and if the letter "K" was detected in the stream, the right button. A target letter appeared with a 40% probability after a minimum of 5 non-target letters. Typically, between 2-3 letter targets appeared during each task event. During the null and baseline intervals, the fixation dot was red, and the letter stream continued but never contained target letters. Observers were instructed to simply maintain fixation during these periods.

Importantly, visual stimulation was identical throughout the experiment; i.e., the letter stream containing targets was presented within every block regardless of task, as were the color patches. This ensured that differences in BOLD signal between attention conditions would not be a result of changes in visual stimulation but rather a result of the amount of attention directed to the stimulus.

One concern with the color patch detection/identification task was the opacity level; we wanted this task to be challenging such that observers would be motivated to continuously attend to the stimulus. Therefore, immediately after getting settled in the scanner (prior to fMRI data collection), observers completed two runs of a staircasing task in which the opacity of the patch was adjusted. The visual stimuli and response key mapping were identical to the main task, with the exception that every block contained the color task. The initial alpha (opacity) value was set to 0.1, and an adaptive staircase (Watson & Pelli, 1983) converged after 40 trials in each run to determine the threshold for 90% accuracy. We took the lower of the two estimates as the observer's individual alpha value for the main task. We also monitored observers' performance on the color task throughout the scan session and adjusted the alpha as needed, to compensate for performance improving over time. On average, the alpha level was 0.053 (± 0.007).

Observers found the color task to be more challenging compared to the fixation task, with mean letter RSVP accuracy 92.79% ± 1.36% and color patch identification accuracy 82.99% ± 1.26%. Across participants, the task performance difference reached statistical significance (paired t-test, $t(9) = 6.9$, $p < 0.001$).

After the completion of the color opacity thresholding and prior to starting the main runs, we also collected two runs of a functional localizer, presented in a block design (14 s on, 14 s off, 182 TRs at 1 s TR, starting and ending with an off block). The localizer stimulus had inner and outer eccentricities identical to the main task stimulus, but was a solid annulus (i.e., no separation into wedges) consisting of a 100% Michelson contrast solid spatial pattern generated by combining radial and spiral gratings, with 10 Hz contrast-reversing flicker. The inner and outer bounds of the localizer were identical to the experimental stimulus (inner diameter 3 dva, outer diameter 17 dva). Data from the functional localizer analysis were later used for voxel selection.

Eye-tracking data acquisition

Eye-tracking data were collected for all observers with a MR-compatible EyeLink 1000 (SR Research, Ontario, Canada) eye tracker system. The sampling rate was 500 Hz. Due to a technical error, all runs from one observer and three runs in another observer were accidentally recorded at 1000Hz and terminated prematurely (after approx. 350 s). For these runs, we cropped the data after 350 s and included the resulting data files in the analysis.

While the behavioral performance suggests that observers were maintaining steady fixation, we also compared eye movements between the two conditions to ensure

there were no significant differences. Blinks were removed (padding 125 ms on either side of a detected blink, resulting in an average of 14.8% ± 2.2% of the eye-tracking data removed per participant), and the absolute distance from fixation was calculated for each attention condition (across orientation differences). The average distance was 0.39 dva (± 0.04) in the fixation task and 0.4 (± 0.04) in the color task, with no significant difference between the two conditions (paired t-test, $t(9) = 0.66$, $p = 0.53$). Moreover, in both conditions the average distance from fixation was well under the radius of the fixation circle (1.5 dva).

MRI data acquisition

All MRI data were collected using a Siemens Prisma 3.0 Tesla scanner (Siemens, Erlagen, Germany) with a 64-channel head coil at the Boston University Cognitive Neuroimaging Center. fMRI data were acquired with simultaneous multi-slice echoplanar T2*-weighted imaging. The fMRI acquisition field of view (FOV) was oriented perpendicular to the calcarine sulcus (FOV = 60 × 112 × 172 mm, TR = 1000 ms, TE = 35 ms, FA = 80°, voxel size = 2 mm isotropic). A whole-brain anatomical scan used to register the functional data was acquired in a separate session (see Population receptive field mapping below) using a T1-weighted multi-echo MPRAGE sequence (FOV = 256 × 256 × 176 mm, 36 slices, TR = 2530 ms, TE = 1.69 ms, FA = 7°, voxel size = 1 mm isotropic).

Population receptive field mapping. Prior to the experimental session, each observer completed a 1.5-2 h population receptive field (pRF) mapping session using standard stimuli and procedures provided by the analyzePRF toolbox (Kay et al., 2013).

pRF mapping results were used to manually draw the boundaries of early visual areas V1, V2, and V3, identified as reversals in polar angle preference.

Anatomical data analyses. The whole-brain T1-weighted anatomical image (voxel size 1 mm isotropic) acquired at the start of the pRF mapping session was analyzed in FreeSurfer (version 5.3) using the recon-all pipeline. The anatomical scan was used to register the functional data with boundary-based registration.

fMRI data analyses and voxel selection. Prior to preprocessing the fMRI data, EPI distortion correction was applied with FSL (Smith et al., 2004) using the reverse phase-encoding method (Andersson et al., 2003). Data were then preprocessed using FS-FAST (Fischl, 2012) with standard motion correction, Siemens slice timing correction, and boundary-based registration (Greve & Fischl, 2009). Following this, we implemented robust rigid registration (Reuter et al., 2010) for voxel-to-voxel alignment between individual runs. No spatial smoothing was applied in preprocessing (FWHM = 0 mm).

Functional localizer data were analyzed using a standard GLM analysis in FreeSurfer in order to identify voxels most responsive to the localizer stimulus. To analyze the main task runs, we used a voxel-wise finite-impulse response (FIR) model in FSL to carry out a univariate analysis of the preprocessed BOLD data. The window size was set to 24 TRs, with a pre-stimulus window of 4 TRs. For each voxel, we thus obtained 24 beta weights for each combination of orientation difference and attentional state. This data was further analyzed with custom MATLAB scripts, and statistical comparisons between attention condition and visual areas were conducted by means of

repeated-measures ANOVA in R. We report variance in our data as standard error of the mean, unless otherwise specified.

Only the top 40% of localizer-responsive voxels were selected for further analysis within each region of interest (ROI), based on their ranking of the localizer GLM significance values. Out of this pool, we excluded any voxels whose pRF eccentricity estimates fell outside the stimulus bounds (1.5 dva inner radius and 8.5 dva outer radius). We also excluded voxels with poor goodness-of-fit of their pRF models (pRF $R^2 < 10\%$). Following this initial voxel selection, the average number of voxels submitted to the fitting procedure (see below) was 476 (SD ± 65) in V1, 343 (± 35) in V2, and 264 (± 39) in V3.

Fitting the orientation-tuned suppression function. In order to obtain the tuned normalization functions for each observer, we first averaged beta weights across all voxels and conditions within each ROI and determined the TR of the peak of this time course. We then took 3 TRs on either side of this peak as the end-points of the averaging window for each voxel for that observer and ROI, resulting in a 7-TR temporal averaging window (equal across conditions). This step aimed to account for differences in the timing of the hemodynamic response between observers and conditions. The resulting functions of beta weights against orientation difference were fit with a half Gaussian function on a voxel-wise basis using the *fmincon* function in MATLAB, as follows:

$$Response = b + Ae^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (2)$$

The fitting procedure was conducted separately for the two attention conditions, and converged on a solution for all voxels. The μ parameter represents the mean of the function and designates where it will be centered, and was fixed to zero. The bandwidth is denoted by the σ parameter, or standard deviation, and was constrained between 8° (the minimum possible difference between the oriented components given the orientation filter bandwidth and minimum orientation difference) and 90°. *b* represents the baseline beta weight (BOLD signal gain) and accounts for the offset of the BOLD response; we left this free parameter unconstrained. *A* is the gain of the Gaussian, and was constrained with an upper bound of 10. Since the mean was set as 0, the fitting procedure largely resulted in an inverted Gaussian (with the lowest response being in the collinear condition), and the gain parameter is negative across voxels for most observers. The gain represents the difference between the fitted response to the orthogonal and collinear conditions. Following parameter estimation, we also calculated R^2 of the Gaussian model fit for each voxel. We found that in each observer and ROI, a subset of voxels had bandwidth estimates that fell along the upper boundary of the constraints for fitting (between 85° and 90°) in at least one condition. A comparison of the R^2 values of the Gaussian fits between voxels whose bandwidth fell above 85° and the remainder revealed significantly lower quality of fits in the high-bandwidth voxels, based on two-factor within-subjects ANOVA with bandwidth group and attention condition as factors ($F(1,9) = 79.92$, $p < 0.001$). In the fixation task, the upper-boundary voxels had R^2 values of 34% (± 5.5%) in V1, 42% (± 4.1%) in V2, and 34% (± 4.7%) in V3, while the remainder of voxels' R^2 values were 60% (± 3.8%) in V1, 65% (± 4.7%) in V2, and 57% (± 6.1%) in

V3. The values were similar in the attend stimulus color task: voxels with higher bandwidth estimates had R^2 values of 35.6% (± 4%) in V1, 39.7% (± 3.6%) in V2, and 37.9% (± 4.3%) in V3, and voxels with lower bandwidths 56.5% (± 4.5%) in V1, 58.5% (± 4.8%) in V2, and 60% (± 4.2%) in V3. We therefore opted for their removal from the dataset. Across observers, we removed 52% ± 2.7% voxels in V1, 45.9% ± 2.8% in V2, and 51.6% ± 1.9% in V3, and the following analyses were conducted using this voxel subset. Prior to voxel exclusion, there was a significant R^2 difference between visual areas in a repeated-measures (RM) ANOVA with factors visual area and attention condition ($F(2,18) = 4.923$, $p = 0.02$), driven by higher R^2 values in V2 as compared to V1 and V3 (V1 vs. V2: $t(19) = -3.18$, $p = 0.015$; V2 vs. V3: $t(19) = 3.15$, $p = 0.016$, after Bonferroni correction); this difference was not observed in the subset of voxels that were kept for further analysis ($F(2,18) = 1.374$, $p = 0.28$). Importantly, in either set of voxels we did not find significant differences in R^2 between the two attention conditions.

Results

Orientation-tuned normalization with and without attention

In line with our previous results, we observed lower BOLD responses for collinear configurations, and larger responses for orthogonal stimuli. For both attentional states, the response gradually increased as a function of increasing orientation similarity, indicating stronger neural suppression as a function of orientation similarity. **Figure 3.2** depicts the observer-averaged BOLD responses (beta weights) as a function of orientation difference under the two attention conditions, for all visual ROIs. While this

general shape can be seen in both tasks, attending to the stimulus evoked a clear overall increase in response to the stimulus.

Voxel-wise tuned suppression strength (defined as orthogonal unattended minus collinear unattended BOLD response) was positive for the majority of voxels for all but one observer **(Figure 3.3B)**. We summarized this measure by taking the median across voxels for each observer/ROI and averaging across observers per ROI. In every ROI, the suppression strength measure was significantly different from zero (one-sample t-test, V1: $t(9) = 2.84$, $p = 0.019$; V2: $t(9) = 4.7$, $p = 0.001$; V3: $t(9) = 4.0$, $p = 0.003$). Overall, suppression strength did not differ between visual areas ($F(2,18) = 2.36$, $p = 0.123$).

Voxel-wise attentional modulation (the difference between beta weights in the attended collinear condition minus the unattended collinear condition) was positive across all observers and ROIs **(Figure 3.3A)**, and significantly different from zero (one-sample t-test, V1: $t(9) = 6.93$, $p < 0.001$; V2: $t(9) = 9.16$, $p < 0.001$; V3: $t(9) = 8.37$, $p < 0.001$), again confirming that visual attention increased the BOLD response. A repeated-measures ANOVA with attention condition and area as factors revealed a significant difference between ROIs ($F(2,18) = 3.95$, $p = 0.038$); however, when following up with Bonferroni-corrected t-tests, we found that none of the inter-ROI differences reached significance. The main effect appeared to be driven by higher attentional modulation in V2 compared to V1 ($t(9) = -2.58$, $p = 0.03$ before correction, $p = 0.089$ after correction).

Quantifying effects of attention. As a means of quantifying changes to the tuned normalization function with attention, we described each voxel's response by fitting with a half-Gaussian function, as done previously (Klímová et al., 2021). The bandwidth of

tuned normalization was denoted by the standard deviation (σ) parameter of the Gaussian. The gain parameter (A) represents the peak of the function. Note that the function is centered on at 0° orientation difference (the mean of the function was set to zero), resulting in an inverted half-Gaussian function where a negative gain represents a lower response for collinear than orthogonal stimulus configurations. We also included an overall response offset parameter (b) to account for the upward shift in responses when the stimuli were attended.

Following the voxel-wise fitting and voxel exclusion (see Methods), The Gaussian parameters: bandwidth, gain, and baseline were submitted to a two-factor within-subjects ANOVA, with the factors of visual area (V1, V2, V3) and attentional state (attend to stimulus or attend to fixation). As can be seen in **Figure 3.4C**, baseline estimates were elevated when observers attended to the stimulus as opposed to attending to the central fixation task; the ANOVA confirmed a significant main effect of attention in this direction ($F(1,9) = 85.17$, $p < 0.001$). The average b estimate (in % signal change) in the fixation condition was 2.21 ± 0.23 in V1, 2.14 ± 0.15 in V2, and 1.64 ± 0.12 in V3, whereas in the color condition the estimates were 2.91 ± 0.25 in V1, 2.99 ± 0.2 in V2, and 2.51 ± 0.18 in V3. There was also a main effect of area ($F(2,18) = 10.71$, $p < 0.001$). Bonferroni-corrected post-hoc tests showed that this effect was driven by significantly higher BOLD response in V1 and V2, compared to V3 (V1 vs V3 $t(19) = 4.66$, $p < 0.001$; V2 vs V3 $t(19) = 10.5$, $p < 0.001$). The interaction term was marginally non-significant ($F(2,18) = 3.488$, $p = 0.053$).

Crucially, the bandwidth of tuned normalization **(Figure 3.4A)** did not differ between attentional states ($F(1,9) = 0.308$, $p = 0.593$) or visual areas ($F(18,2) = 0.414$, $p = 0.667$), with a non-significant interaction term ($F(2,18) = 0.004$, $p = 0.996$), indicating lack of modulation in orientation-difference tuning as a result of attending to the stimuli. Likewise, changes in the gain parameter **(Figure 3.4B)** were also non-significant, both between attention conditions ($F(1,9) = 0.018$, $p = 0.897$) and between visual areas ($F(2,18) = 2.404$, $p = 0.119$), with no interaction ($F(2,18) = 0.534$, $p = 0.595$). In the unattended condition, the average bandwidth was 20.6° (± 1.35°) in V1, 20.8° (± 0.84°) in V2, and 20.23° (± 0.94°) in V3, while in the attended condition, bandwidths were 21.96° (± 2.35°) in V1, 22.06° (± 1.85°) in V2, and 21.5° (± 1.93°) in V3.

In order to verify that our voxel exclusion criteria based on the bandwidth estimates did not bias the overall statistical results, we repeated the above tests having skipped the voxel exclusion step. The results were consistent between the two voxel subsets; the only exception besides the R^2 measures detailed in Methods is the interaction term between visual area and attention condition for the offset (b) parameter, which shifted from marginally non-significant ($F(2,18) = 3.488$, $p = 0.053$; see above) to marginally significant when all initially fitted voxels were submitted to the ANOVA ($F(2,18) = 3.689$, $p = 0.0454$).

Discussion

Our data show overall significantly increased BOLD signal in the visuocortical responses to oriented stimuli when stimuli are attended to, compared with attention withdrawn away. Importantly, the absence of significant shifts in the bandwidth and gain

of orientation-tuned suppression further point to a lack of qualitative change in the response magnitude to varying feature differences, suggesting that attending to a visual stimulus did not alter the sensitivity to orientation differences in early visual areas.

We designed our attention manipulation to ensure participants were directing covert, endogenous attention to the entire oriented stimulus (by virtue of unpredictability of the target location), and to avoid any exogenous attention or pop-out effects (through the use of individual opacity thresholds for each observer and constant monitoring of performance). Covert visual attention, and its effects on sensory gain, has most frequently been studied in the context of how it affects neural responses to varying stimulus intensity (Carrasco, 2011; Itthipuripat et al., 2019; Martínez-Trujillo & Treue, 2002; Reynolds & Heeger, 2009). A common finding in electrophysiological experiments is a multiplicative gain increase of the neural contrast response (McAdams & Maunsell, 1999; Williford & Maunsell, 2006). In contrast, in the fMRI literature, a majority of findings report an additive increase in attended states over unattended (Buracas & Boynton, 2007; Itthipuripat et al., 2019; Murray, 2008). To explain this discrepancy, it has been suggested that additive baseline effects seen in fMRI could simply be the result of insensitivity of the BOLD signal to modifications of stimulus-driven activity (Itthipuripat et al., 2014). A recent study from our lab presented evidence that attention in fact produces non-additive effects in contrast response functions (CRF) measured with BOLD fMRI (Foster & Ling, 2022). It is important to note that in this study, we did not manipulate stimulus intensity, and instead measured responses to orientation offsets at a fixed 50% contrast level. Therefore, the increase in our orientation-tuned normalization

baseline parameter should not be confused with the baseline shifts sometimes found in attention studies that measure the CRF (Buracas & Boynton, 2007; Itthipuripat et al., 2019; Murray, 2008). Our results could be driven by either multiplicative or additive modulation in the domain of contrast; within the confines of recording at a single contrast level, we can only comment on attentional modulation of the orientation-tuned suppression function. Similarly, we are also not measuring orientation tuning, but rather tuning to the orientation differences in our stimulus, a higher-order statistic.

Our findings square with theories of efficient coding, which posit that sensory neurons are adapted to the signals they are most exposed to, and in particular, that signal processing in the visual cortex is optimized to the prevalent image statistics we find in our natural environments (Coen-Cagli et al., 2015; Simoncelli & Olshausen, 2001; Vinje & Gallant, 2000). For instance, the estimated orientation-tuned suppression bandwidths in this study are comparable with the parameter estimates from our previous work, in which average bandwidth estimates were between 23.1° (V1), 24.4° (V2), and 22.8° (V3) (Klímová et al., 2021), and the pattern of suppression strength as a function of orientation similarity bears resemblance to the prevalence of orientation differences in nearby regions of a scene (Sigman et al., 2001). While it has been established that attention alters subjective appearance in ways that enhance the selective processing of attended areas, objects, or features (Carrasco et al., 2004; Carrasco & Barbot, 2019), a change in tuned normalization bandwidth would imply a change in the processing of second-order statistics (i.e., a change in the subjective salience of feature *differences*), which is to some degree dependent upon our visual system's apparent adaptation to the typical

composition of orientations and other basic visual features in our environments (Simoncelli & Olshausen, 2001).

Figures

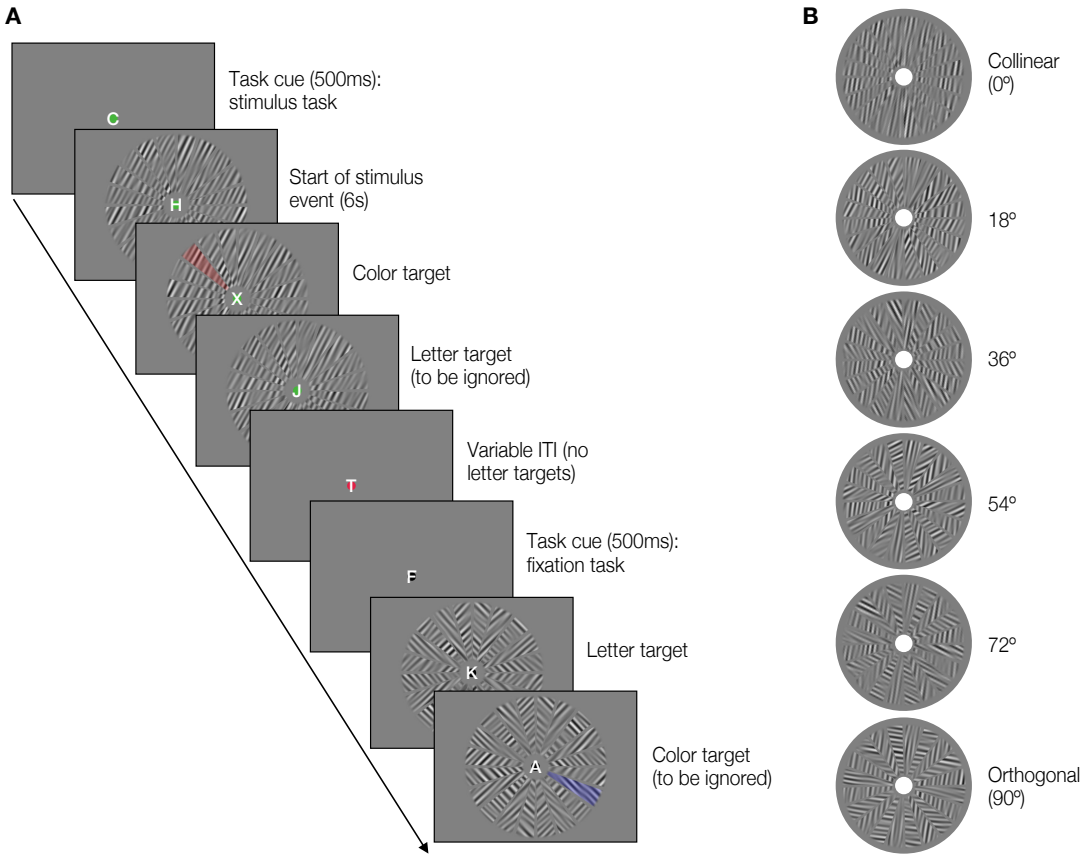

Figure 3.1 Experimental stimuli and procedure.

(A) Illustration of experimental procedure, depicting a sample of two stimulus event blocks, with a blank interval in between. Here, the observer is shown a color task cue, followed by a stimulus event with a collinear condition. After an inter-trial interval with no stimulus, a fixation task cue was shown, followed by a fixation task block with an orthogonal condition. Note that the only visual change on the screen between the two conditions is the fixation dot color. In this depiction, the opacity of the color patches has been increased for visibility, as has the size of the fixation dot and the letter. The larger white fixation circle is not depicted in this illustration. (B) Experimental stimuli. Stimuli are all depicted with a starting orientation of 9° from vertical. For illustration purposes, these stimuli are rendered at a lower spatial frequency and a higher contrast.

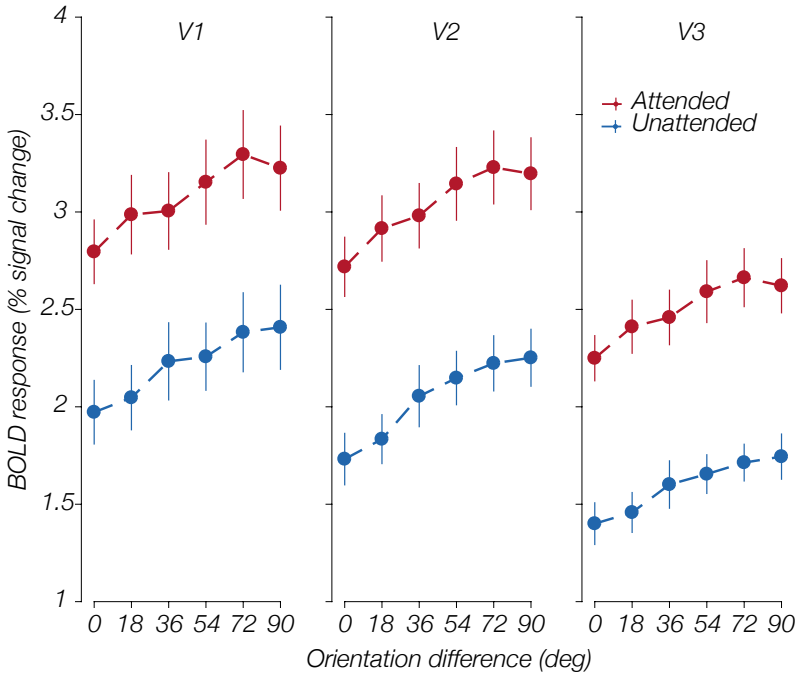

Figure 3.2 Orientation-tuned suppression of the BOLD response.

Observer-averaged (n = 10) BOLD response (± 1 SEM) as a function of stimulus orientation difference in the attended (color task) and unattended (fixation/letter task) conditions.

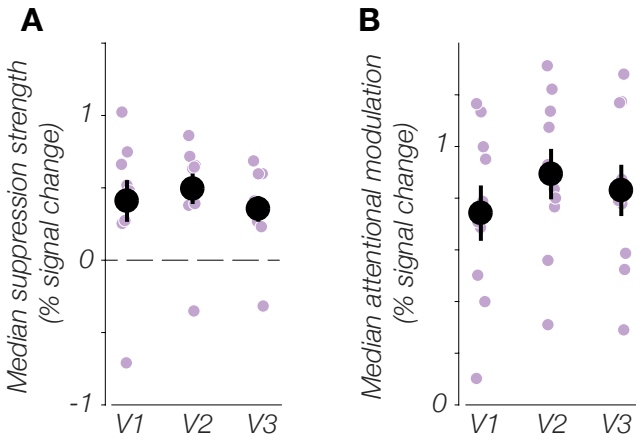

Figure 3.3 Suppression strength and attentional modulation.

Observer-averaged (n = 10) median (A) suppression strength, defined as voxel-wise difference between beta weight in orthogonal minus collinear unattended conditions, and (B) attentional modulation, computed for each voxel as the difference between beta weight in collinear attended condition minus collinear unattended condition. Error bars represent ± 1 SEM.

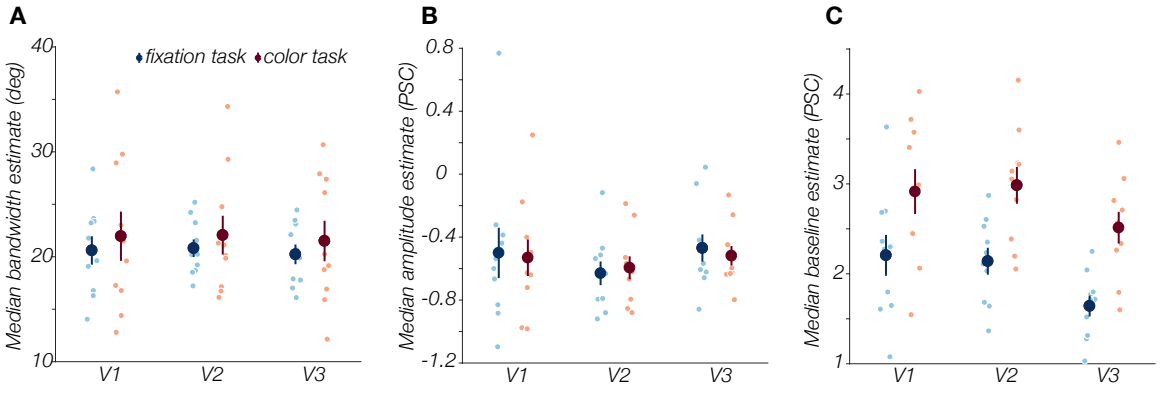

Figure 3.4 Gaussian parameters under the two attentional states.

Observer-averaged (n = 10) Gaussian parameters describing the tuned normalization functions. (A) bandwidth, (B) gain, (C) baseline. Error bars represent ± 1 SEM. PSC = % signal change.

CHAPTER FOUR: HOW DOES TUNED NORMALIZATION MODULATE POPULATION CONTRAST RESPONSE FUNCTIONS?

Introduction

Our perception of any stimulus is heavily influenced by its spatial context. For instance, the apparent contrast of a central grating stimulus can be lowered by the presence of a high-contrast surrounding annulus (Cannon & Fullenkamp, 1991; Xing & Heeger, 2000; Zenger-Landolt & Heeger, 2003); this perceptual phenomenon is mirrored in the suppression of visuocortical neural responses to central receptive field (RF) stimulation when a surround is added (Cavanaugh et al., 2002a; DeAngelis et al., 1994; Webb et al., 2005).

Surround modulation in the visual cortex has been extensively studied in animal electrophysiology (Carandini et al., 1997; Cavanaugh et al., 2002a; DeAngelis et al., 1994; Polat et al., 1998; Sengpiel et al., 1998; Webb et al., 2003, 2005) as well as human psychophysics and neuroimaging (Cannon & Fullenkamp, 1991; Petrov et al., 2005; Xing & Heeger, 2000), and is presumed to be a mechanism facilitating the use of spatial context to parse visual scenes (Angelucci et al., 2017; Schwartz & Simoncelli, 2001). The addition of a surround stimulus is typically found to be suppressive (Cavanaugh et al., 2002a; DeAngelis et al., 1994; Sengpiel et al., 1998; Webb et al., 2003, 2005), although in some configurations it can be facilitatory (Ichida et al., 2007; Schallmo et al., 2018; Sengpiel et al., 1998; Toth et al., 1996).

The properties of the center and surround stimuli dictate whether, and how much, suppression will be present. Crucial factors are the contrast and relative orientations of

the two components. Typically, higher surround contrasts evoke stronger suppression, and the strongest suppression occurs when the surround and central stimuli match in orientation and spatial frequency (Cavanaugh et al., 2002a; Self et al., 2014; Webb et al., 2005). The influence of the surround on the center is widely accounted for as a form of divisive modulation (Carandini & Heeger, 2012; Cavanaugh et al., 2002a; Fang et al., 2021), in which the excitatory drive from the center stimulus is divided by a proportional suppressive drive comprised of a more broadly tuned pool of units responding to both the center and the surrounding region of space, such that higher drive from the center causes stronger suppression. The divisive normalization model (Carandini & Heeger, 2012; Heeger, 1992) has been put forth as an explanatory account of nonlinear behavior of early visual cortical neurons, such as the saturating contrast response function and cross-orientation suppression, but also surround suppression (Carandini & Heeger, 2012). However, the divisive normalization model has been critiqued for its inability to account for surround facilitation, and does not explicitly address the anatomical substrates giving rise to surround modulations (Ichida et al., 2007; Somers et al., 1998). Another class of model are recurrent circuit models, which are based in the known anatomy of V1 cortical circuits, and incorporate bottom-up, feedforward inputs as well as lateral horizontal connections between neurons in a visual area, and sometimes also feedback connections from higher visual areas (Angelucci et al., 2017; Schwabe et al., 2006; Somers et al., 1998). In circuit models, surround modulation comes primarily from horizontal and/or feedback connections, and the sign (facilitation vs. suppression) depends on the relative

activation of excitatory and inhibitory inputs. However, both classes of model assume that non-linear, divisive computations give rise to surround suppression.

Surround modulation can be characterized by measuring the neural or psychophysical contrast response function (CRF) to a central stimulus and observing how it rescales upon introducing a surround stimulus. In the literature, the CRF is typically described by the Naka-Rushton equation (Albrecht & Hamilton, 1982; Naka & Rushton, 1966). A rightward shift, or contrast gain, of the CRF as a result of placement of a high-contrast surround, manifests as an increase in the semi-saturation parameter, C_{50}. Similar to contrast adaptation, this serves to shift the most sensitive portion of the neuron's response curve towards the predominant contrast level in the environment, increasing its sensitivity (DeAngelis et al., 1994; Ohzawa et al., 1985). Response gain, or a downward shift most prominent at high contrasts, is described as a decrease in the R_{max} parameter and suggests decreased responsivity. V1 electrophysiological investigations of CRF surround modulation have shown considerable heterogeneity among the recorded units. The primary finding in some studies has been contrast gain (Carandini et al., 1997; DeAngelis et al., 1994; Webb et al., 2005), suggesting that the effect of increasing surround contrast is similar to that of contrast adaptation, shifting the most sensitive portion of the CRF towards the predominant contrast in the environment. Other experiments find primarily response gain (Webb et al., 2003). Overall, most studies report a mixture of effects (Cavanaugh et al., 2002a; Sengpiel et al., 1998), where a larger proportion of neurons showed response gain but some were better fit with a contrast gain modulation. However, all above cited studies find greatest suppression with collinear

surrounds, and weaker suppression with orthogonal surrounds, supporting the notion that divisive normalization is sensitive to stimulus features.

In human psychophysics, the apparent contrast of a central stimulus is lower in the presence of a high-contrast surround (Cannon & Fullenkamp, 1991; Xing & Heeger, 2000; Zenger-Landolt & Heeger, 2003), and this suppressive effect is again stronger with collinear surrounds as compared to orthogonal (Cannon & Fullenkamp, 1991). Similarly, in neural population measures such as fMRI, the BOLD response to a stimulus is decreased by the simultaneous presentation of a surround (Kastner, 1998; Zenger-Landolt & Heeger, 2003), with collinear stimuli inducing more suppression than orthogonal (McDonald et al., 2009; Pihlaja et al., 2008; Williams et al., 2003). However, a full CRF is rarely measured in human fMRI studies. A major obstacle to the study of CRF modulation by surround suppression in human visual cortex in this modality has been the difficulty in obtaining non-linear CRFs in this way (Boynton et al., 1999; Itthipuripat et al., 2019; Tootell et al., 1998). Recently, work from our group showed that non-linear CRFs can be measured in fMRI with the use of an adaptation paradigm (Foster & Ling, 2022; Vinke et al., 2022), where an extended period (60 s) of adaptation to a low contrast value is combined with top-up adaptation throughout the experiment. Adaptation lowers the heterogeneity of voxel responses (Vinke et al., 2022), which allows us to capture the underlying non-linear responses. In this study, we modify the adaptation paradigm to examine the modulations of the CRF in human visual cortex specific to the relative orientations of the center and surround stimuli. We measure the CRF to a central grating stimulus at 9 contrast levels, while the center is surrounded by a large 100% contrast

annulus grating either collinear or orthogonal with the center. We find that while the CRFs of voxels near fovea largely overlap, the CRFs of the voxels adjacent to the center-surround boundary primarily exhibit contrast gain in response to the collinear surround relative to orthogonal, supported by a rightward shift of the CRF and an increase in the semi-saturation parameter. However, we also find considerable variability in the voxel-wise CRFs.

Methods

Observers

Ten observers (8 female) took part in the experiment. All were between the ages of 18-35 and reported normal or corrected-to-normal visual acuity. All participants gave their written informed consent, and the study was approved by the Boston University Institutional Review Board. Observers received monetary compensation for their participation, except one (one of the authors of the study).

Apparatus and Stimuli

Stimuli were programmed and rendered on a MacBook Pro (OS X 10.7) using MATLAB (2015b; Mathworks, Natick, MA) and Psychophysics Toolbox (Brainard, 1997). The stimuli were displayed on a rear-projection screen in the scanner bore, using a gamma-corrected projector (ProPixx DLP LED, VPixx Technologies) and observers viewed them via a front-surface mirror affixed to the head coil. Participants were provided with hearing protection and a 2-button box for behavioral responses.

The visual stimulus was a 2 cycles/degree (cpd) center grating (inner radius 0.75 dva from central fixation, outer radius 2.95 dva), which varied in contrast throughout each fMRI run, surrounded by a 2 cpd annular grating (inner radius 3.05 dva, outer radius 8.5 dva), with a 0.1 dva gap between the central and surround component (**Figure 4.1A**). The small gap was chosen based on prior work which found strong center-surround interactions with minimal spatial separation between the two components (Cannon & Fullenkamp, 1991; Petrov et al., 2005; Phillips et al., 2021; Poltoratski et al., 2019). Prior work also informed the eccentric location of the center-surround boundary; surround suppression tends to be stronger when stimuli are presented away from fovea (Xing & Heeger, 2000), and saturating nonlinearities were observed with the adaptation paradigm predominantly in the foveal and parafoveal region, with saturation declining with eccentricity (Vinke et al., 2022). Both gratings were embedded in a Gaussian envelope. The contrast of the center grating varied over nine logarithmically spaced contrast levels (2.67%, 4.0%, 5.33%, 8.0%, 16%, 32%, 48%, 64%, 96% Michelson contrast), while the surround grating contrast was always 100%. Both gratings had their spatial phase updated every 100 ms to a randomly chosen value, independently of each other. The surround grating could either be collinear or orthogonal with respect to the center. The central grating orientation remained identical throughout each run, and was either 45° or 135° in alternating runs, with starting orientation counterbalanced between observers. Stimuli were presented on a mean luminance background.

MRI data acquisition

All MRI data were collected at the Center for Cognitive Neuroimaging center at Boston University on a Siemens 3T Prisma scanner with a 64-channel head coil. In a single two-hour session. fMRI data were acquired with simultaneous multi-slice (multi-band acceleration factor 5) echoplanar T2*-weighted sequence (voxel size $2mm^3$, TR = 1,000 ms, TE = 30 ms, flip angle = 64°, FOV = 208 × 208 × 140 mm). Prior to this session, each participant also went through a separate population receptive field (pRF) mapping session using the same T2*-weighted protocol, in addition to a high-resolution anatomical scan (T1-weighted multi-echo MPRAGE sequence, FOV = 256 × 256 × 176 mm, 36 slices, TR = 2530 ms, TE = 1.69 ms, FA = 7°, voxel size = $1mm^3$).

Experimental procedure

Main task. The main task had 498 TRs, and most participants completed 10 runs (one completed 8, and two completed 9). Stimuli were presented in an event-related design, with 4 s event duration and jittered inter-trial interval between 6-17 s. The event schedules were generated using the FreeSurfer tool Optseq2 (Dale, 1999). To promote nonlinear contrast response functions, we used a contrast adaptation paradigm previously established in our lab (Foster & Ling, 2022; Vinke et al., 2022). Following a 4 s baseline period with a mean luminance screen, the phase-jittered central grating was presented for 60 s at 16% contrast (adapting contrast) in an initial adaptation block. Following adaptation, the event-related stimulus presentation began. During the stimulus event, the center grating changed contrast to the target contrast for that event, and was surrounded by either a collinear or orthogonal 100% contrast grating. The inter-trial intervals served

as top-up adaptation periods, during which the center grating again changed contrast to the adapting contrast. An example stimulus sequence is depicted in **Figure 4.1B**. Each of the 9 center contrast levels (including the adapting contrast) was presented four times within an fMRI run, twice with a collinear surround and twice with an orthogonal surround.

Participants were engaged in a rapid letter detection and identification task at fixation. The small (0.1 dva) fixation dot in the center of the screen was red, and surrounded by a white circular 1.5 dva diameter annulus. White letters were displayed within this annulus, in front of the fixation point, continuously throughout the run. Participants' task was to monitor this letter stream for letters 'J' and 'K' amid 10 other distractor letters ('X', 'L', 'V', 'H', 'S', 'A', 'C', 'P', 'Z', 'Y'). A new letter was presented every 200 ms, and participants were asked to press the left button on the response box as soon as they detected 'J', and the right button for 'K'. At the end of each run, performance accuracy was displayed to the participants for feedback. Accuracy across participants was 90.2% on average (± 2.4% SEM).

Functional localizer. Each session began with two runs of a functional localizer, intended to isolate voxels responding to the center and the surround stimulus areas of the visual field. The localizer had a stimulus on – stimulus off blocked design, with 208 TRs (1 s TR), and 16 s blocks, beginning and ending with an off block. The localizer stimulus was a 100% Michelson contrast 2 cpd achromatic checkerboard with the same inner and outer diameter as the main stimulus, on a mean luminance background, and the

behavioral task was identical to the main experiment. Following the localizer runs, participants began the main task.

Population receptive field mapping session. For each observer, pRF mapping was carried out in a separate session, as described in the Methods sections of the previous two chapters, using stimuli and analysis code from the analyzePRF toolbox (Kay et al., 2013). The results were used to draw cortical surface labels outlining early visual areas V1, V2, and V3, by identifying polar angle preference reversals. The early visual area labels then served as a tool in voxel selection for functional data analysis.

MRI data analyses

Anatomical data. The 1 mm^3 T1 anatomical images acquired during the pRF mapping session were analyzed in FreeSurfer using the recon-all pipeline. The results were used to register the functional data to the anatomical data.

fMRI preprocessing and beta weight estimation. Reverse-phase encoding (Andersson et al., 2003) was used to correct EPI distortion in the functional data in FSL (Smith et al., 2004). Data were preprocessed with FS-FAST (Fischl, 2012) with no spatial smoothing (FWHM = 0 mm), implementing standard motion correction, Siemens slice timing correction, and boundary-based registration (Greve & Fischl, 2009). We used robust rigid registration (Reuter et al., 2010) to achieve accurate voxel-to-voxel correspondence between functional runs within a session, aligning the middle TR of each run to the middle TR of the first run of the session. To identify voxels responsive to the stimuli, the functional localizer data for each localizer type (center and surround) were analyzed in FreeSurfer with a GLM analysis following robust registration. The main task

data were further processed using custom MATLAB scripts. We extracted voxels that fell within the pRF labels V1, V2, and V3. Following the removal of the beginning 64 TRs from each run (the 4 s initial baseline + the 60 s initial adaptation period), the time series data were low-pass filtered (filter cutoff 0.01 Hz), converted to % signal change by dividing the BOLD signal at each time point by the average BOLD signal value of the run, and concatenated.

We constrained our voxel inclusion as follows: first we selected only voxels responding to either the center or surround localizer, defined as a GLM *p*-value of 0.05 or less. Out of these voxels, we further selected only those with a pRF goodness-of-fit (R^2) of 10% or above, and also those whose eccentricity estimates fell within the stimulus bounds (i.e., between 0.75 and 8.5 dva). Furthermore, we ensured that voxels whose labels overlapped were removed. After the application of these criteria, we had on average 719 ± 174 (SD) voxels in V1, 485 ± 88 in V2, and 335 ± 42 in V3.

Contrast response estimation. After finalizing the initial voxel selection, we implemented a voxel-wise finite impulse response (FIR) analysis in MATLAB, with a window size of 20 s following stimulus onset, resulting in 20 beta weight estimates for each condition (center-surround orientation offset and center contrast level). Finally, we computed the mean beta weight in each condition within an averaging window of 4 – 8 TRs after stimulus onset, accounting for the hemodynamic response delay and capturing the peak of the hemodynamic response function for each observer and condition, resulting in a voxel-wise contrast response function of 9 points (contrast levels) per condition.

Contrast response function model fitting. The contrast response function was then fit with a variant of the Naka-Rushton equation (Albrecht & Hamilton, 1982; Naka & Rushton, 1966):

$$R(c) = R_{max}\frac{c^n}{c^n + {C_{50}}^n} + b \quad (3)$$

Here, the BOLD response (R) at each contrast level (c) is determined by the maximum attainable response (R_{max}), the contrast at the semi-saturation point (the semi-saturation constant, C_{50}), an exponent (n), and an additive baseline parameter (b). MATLAB's *fmincon* function was used to implement the fit. We constrained the R_{max} parameter to be between 0 and 10 (beta weight, or % signal change), and the C_{50} parameter to be between 1 and 80 (% contrast). The baseline parameter was fixed per voxel to the average of the voxel's responses to the lowest contrast between the collinear and the orthogonal surround condition. Furthermore, we did not anticipate significant changes in the n parameter based on existing literature (Carandini et al., 1997; Cavanaugh et al., 2002a; DeAngelis et al., 1994; Webb et al., 2005); therefore, we opted to fix the value of n to 2 in each voxel (Carandini & Heeger, 2012; Tolhurst & Heeger, 1997). The fitting procedure converged on a solution for all voxels. A goodness-of-fit estimate was obtained by computing the R^2 of the Naka-Rushton fit for each voxel. We removed voxels where the Naka-Rushton R^2 was less than 0 in either condition (40.7 ± 5.1% SD in V1, 49 ± 6.8% in V2, and 47 ± 10.9% in V3). Model fitting was conducted in MATLAB, while most statistical tests were performed in R.

Eye position monitoring

Throughout the experimental session, participants' gaze was monitored using an MR-compatible eye-tracking setup (EyeLink 1000, SR Research, Ontario, Canada) with a sampling rate of 1,000 Hz (3 observers) or 500 Hz (7 observers). After excluding blinks, the average eye deviation from the fixation point in the center of the screen across participants was 0.21 dva ± 0.09 dva SE in horizontally and 0.2 dva ± 0.11 dva SE vertically. This is well within the bounds of the fixation circle, whose radius was 0.75 dva. Therefore, participants maintained reliable fixation throughout the experimental session.

Results

Contrast response functions under orientation-tuned suppression

Given the spatial layout of our full-field stimulus, we reasoned that any orientation-tuned modulation would be most apparent for voxels whose pRF location (eccentricity) is near the center-surround boundary. Instead of averaging the voxel-wise CRFs across the whole ROI, we binned the voxels into 8 bins based on their pRF eccentricity. We first divided the stimulus into two portions: center (between 0.75 and 3.05 dva radius) and surround (between 3.05 and 8.50 dva radius), with the inner radius of the surround stimulus serving as the dividing line. We then divided each half of the display into four equal-sized eccentricity bins. As depicted in **Figure 4.2**, there is a gradual gain modulation in the central portion of the stimulus; in the bin closest to fixation, the contrast responses to the collinear and orthogonal flanked condition are largely overlapping. After this, the two begin to diverge, and in the fourth bin (the center

stimulus band abutting the surround), the gain modulation by orientation-tuned suppression appears strongest, with signs of stronger suppression in the collinear condition. Differences between collinear and orthogonal condition persist in the first surround bin, but the outermost bins show flatter responses for both conditions, which again largely overlap. Differences between the surround configurations are summarized in **Figure 4.3**.

Naka-Rushton parameters

In order to compare the voxel-wise Naka-Rushton parameters between the two surround configurations, we selected voxels based on whether the extent of their pRFs included the boundary between the center and the surround. For each voxel, we added and subtracted the pRF size estimate (which represents the standard deviation of the 2-dimensional Gaussian used to model the pRF (Kay et al., 2013)) with the voxel's pRF eccentricity estimate, to obtain the full extent of the pRF on the stimulus in dva. Across participants, on average 109 (± 48 SD) voxels in V1, 51 (± 14) voxels in V2, and 43 (± 10) voxels in V3 fulfilled this criterion. The average eccentricity of the voxels was 3.18° from fixation across ROIs (± 0.71° SD), ranging between 1.41° and 8.02°.

In this subset of voxels, we calculated median C_{50} and R_{max} estimates per observer and ROI, and compared the parameters in each ROI between the two surround conditions using a one-sided pairwise Wilcoxon test (reflecting our reasoning that if suppression is stronger in the collinear configuration, we should see a higher C_{50} in this condition, and/or a lower R_{max}, as seen in electrophysiology). The results are shown in **Figure 4.5**. C_{50} was overall higher in the collinear condition: in V1, observer-averaged median C_{50}

was 50.89 ± 7.5% contrast (SEM) for the collinear configuration and 38.37 ± 6.66% contrast in the orthogonal configuration, in V2, collinear C_{50} was 34.55 ± 8.78% contrast while orthogonal was 10.75 ± 1.2% contrast, and in V3, collinear C_{50} was 32.23 ± 7.59% contrast and orthogonal 20.23 ± 5.99% contrast; however, this difference reached significance only in V2 (Z = -2.25, p = 0.012, after Bonferroni correction to the number of ROIs). R_{max} did not differ between conditions in any ROI. The observer-averaged median collinear R_{max} was 1.27 ± 0.08% signal change (SEM) and orthogonal was 1.27 ± 0.09% signal change in V1; in V2, R_{max} was 0.54 ± 0.05% signal change in the collinear condition and 0.64 ± 0.1% signal change in the orthogonal condition, and in V3, 0.47 ± 0.04% signal change while orthogonal R_{max} was 0.52 ± 0.05% signal change.

For a more comprehensive overview of surround orientation effects, linear mixed-effects models were used, incorporating all voxels in the center-surround boundary selection and including observer as a random effect (obtaining a model intercept for each observer), to model C_{50} and R_{max} (in separate models) as predicted by surround orientation and ROI. In predicting the values of the semi-saturation parameter, the effect of surround orientation was statistically significant, confirming lower C_{50} in the orthogonal condition (*beta* = -8.34, 95% *CI* [-10.12, -6.57], *t(4046)* = -9.19, $p < 0.001$). Similarly, the effect of ROI was significant, with overall lower C_{50} in V2 (*beta* = -14.82, 95% *CI* [-17.02, -12.61], *t(4046)* = -13.18, $p < 0.001$) and V3 (*beta* = -10.82, 95% *CI* [-13.15, -8.48], *t(4046)* = -9.07, $p < 0.001$) compared to V1. For R_{max}, the effect of surround orientation was not statistically significant (*beta* = 0.06, 95% *CI* [-0.03, 0.15], *t(4046)* = 1.33, p = 0.182), and the effects of ROI was again significant, reflecting lower

R_{max} in V2 (*beta* = -0.81, 95% *CI* [-0.91, -0.70], *t(4046)* = -14.57, $p < 0.001$) and V3 (*beta* = -1.12, 95% *CI* [-1.24, -1.01], *t(4046)* = -19.14, $p < 0.001$) compared to V1. However, the model fits, and therefore, explanatory power, were only weak to moderate (R^2 for the C_{50} model was 0.13 and for the R_{max} model, 0.14). We discuss the probable reasons for this below.

Voxel-wise variability and non-saturation. The voxel-wise CRF, and hence, Naka-Rushton parameter variability, was substantial in all three visual areas. Scatterplots in **Figure 4.4** give a picture of voxel variability and voxel-wise relationships, for each Naka-Rushton free parameter, between the collinear and orthogonal configuration. **Figure 4.6** shows several illustrative individual voxels. Importantly, we find that while most voxels show non-linear CRFs, many of them failed to saturate at high contrast. Out of the total of 2,026 voxels fulfilling our center-surround boundary criteria, 865 did not show saturation in at least one condition; we defined this as a C_{50} value above 75% contrast, to reflect the fact that most of these voxels showed a C_{50} very close to the fitting algorithm upper constraint of 80%, suggesting that the true best-fitting C_{50} would have been outside of the range of possible contrasts were the fitting not constrained by an upper boundary. The lack of saturation in these voxels renders the estimates of the R_{max} and C_{50} parameters less meaningful, limiting our ability to capture reliable differences between the two conditions, despite the visible % signal change differences seen in **Figures 4.2 and 4.3**.

Collinear facilitation

Aside from CRF modulations, we were also interested in testing for the presence of collinear surround facilitation effects after observing slightly higher average BOLD signal in the collinear surround condition compared to orthogonal at the lowest center contrast level, particularly in the stimulus segment closest to fixation (see **Figure 4.2**, leftmost panel). However, paired samples t-tests in each ROI comparing the mean % signal change values in the collinear and orthogonal conditions at this contrast level determined that these differences did not reach statistical significance (V1: $t(9) = -2.07$, $p = 0.067$; V2: $t(9) = -1.02$, $p = 0.34$; V3: $t(9) = -0.95$, $p = 0.37$).

Discussion

We investigated how the gain of contrast responses is modulated by orientation-tuned suppression in the early human visual cortex under collinear and orthogonal high-contrast surround. Visible changes in modulation of the CRF were observed predominantly in voxels whose pRF location and size positioned them such that they received stimulation from both the center and surround stimuli, and was maximal in center voxels directly bordering the surround annulus. This contrasts with the strong overlap between the two CRFs in the near-fixation voxels in the center stimulus. Due to low proportions of voxels showing a fully saturating CRF, we were limited in our ability to reliably fit the Naka-Rushton model on a voxel-wise level, and therefore the corresponding statistical comparisons of the semi-saturation constant and the maximal response parameters are also not a full characterization of the CRF changes we observe in the BOLD data. However, the trends in the data strongly suggest a higher semi-saturation

constant in the collinear surround condition as compared to orthogonal surround, indicating a rightward shift of the CRF as a result of stronger suppression caused by the collinear annulus. The R_{max} values remained unchanged between the two surround configurations, suggesting that the orientation-tuned influence of surround suppression on average causes a contrast gain rather than response gain.

The lower rate of saturation in our data diverges somewhat from other studies utilizing adaptation to recover saturating nonlinearities in the CRF (Foster & Ling, 2022; Vinke et al., 2022). Likely contributions are the relative lack of stimulus optimality for early visual cortex, specifically, we did not account for cortical magnification in the stimulus spatial frequency, which was done by Vinke et al. (2022), and the center grating stimulus was not oriented radially from fixation, as done in Vinke et al. (2022), and to some extent in Foster and Ling (2022). Our stimulus was intended to maximize the perceptual suppression from the high-contrast surround presentation, which necessitated less regard for stimulus optimality than previous work. However; despite this, we were able to determine an overall trend towards contrast gain as a result of collinear surround modulation, compared to orthogonal. A possible direction for future analysis would be to use pRF information to isolate the voxel time-courses for each voxel when its preferred radial orientation was present in the center, as the radial optimality may contribute to saturation (Vinke et al., 2022), which could, in turn, facilitate better voxel-wise fits to the Naka-Rushton.

Our results are in agreement with prior fMRI studies in early visual areas demonstrating the orientation dependency of surround suppression, in which parallel

surrounds induced stronger BOLD signal suppression compared to orthogonal surrounds (McDonald et al., 2009; Schallmo et al., 2016; Williams et al., 2003). Likewise, a number of psychophysical investigations have shown stronger perceptual suppression of a target grating by a parallel high-contrast annulus compared to orthogonal (Cannon & Fullenkamp, 1991; Solomon et al., 1993; Xing & Heeger, 2000). Mirroring prior electrophysiological work, we see considerable variability among the individual CRF measurements. There are multiple V1 recording studies which observed primarily contrast gain (Carandini et al., 1997; DeAngelis et al., 1994; Webb et al., 2005), as well as others which find large variability in surround suppression (Cavanaugh et al., 2002a; Sengpiel et al., 1998).

When it comes to surround facilitation, we observed higher BOLD responses in the collinear surround compared to orthogonal at the lowest center contrast, but this effect was confined to the innermost voxel bin (see **Figure 4.2**) and not statistically significant. A few neuroimaging and psychophysical studies report surround facilitation (Kapadia et al., 1995; Polat & Norcia, 1996; Xing & Heeger, 2001), although suppression is a more common finding (Cavanaugh et al., 2002a; DeAngelis et al., 1994; Sengpiel et al., 1998). Prior research has shown that both suppression and facilitation are dependent on relative contrasts and orientations of the center and the surround, but also on the size of the surround, and that increasing the size of the surround causes suppression, rather than facilitation (Cannon & Fullenkamp, 1993; Xing & Heeger, 2001). Furthermore, it appears that the presence of a gap between the center and surround affects whether or not facilitation is observed. Facilitation was primarily observed in studies where the

immediate area surrounding the center did not receive stimulation (Ichida et al., 2007; Polat et al., 1998; Polat & Norcia, 1996; Shushruth et al., 2012); it has been suggested that stimulating the near-surround with a high-contrast stimulus may mask the facilitatory interactions (Ichida et al., 2007). Here we only tested a large surround in close proximity to the center, conditions which in recording studies are more likely to cause suppression over facilitation (Henry et al., 2013).

Our results are generally in line with the proposed role of surround suppression in early visual cortex in the context of scene processing. The rightward shift of the CRF in response to a high-contrast annulus is thought to serve a similar purpose as the adaptation-induced CRF shift in the temporal domain (Müller et al., 2003), whose role is to bring the most sensitive portion of the CRF towards the ambient contrast level in the region around the RF (Ohzawa et al., 1985), optimizing the sensitivity of the neuron through divisive computations. More broadly, a suggested purpose of feature-dependent surround suppression is to serve texture segmentation (Angelucci et al., 2017; Knierim & Van Essen, 1992; Levitt & Lund, 1997; Müller et al., 2003; Schwartz & Simoncelli, 2001), and both differences in contrast and orientation signal the presence of areas of higher interest in a visual scene possibly containing borders between objects or textures. Therefore, suppressing signals from similar regions and enhancing signals from bordering regions with very different textures might achieve higher efficiency in transmitting information via visuocortical spikes (Levitt & Lund, 1997; Müller et al., 2003).

Figures

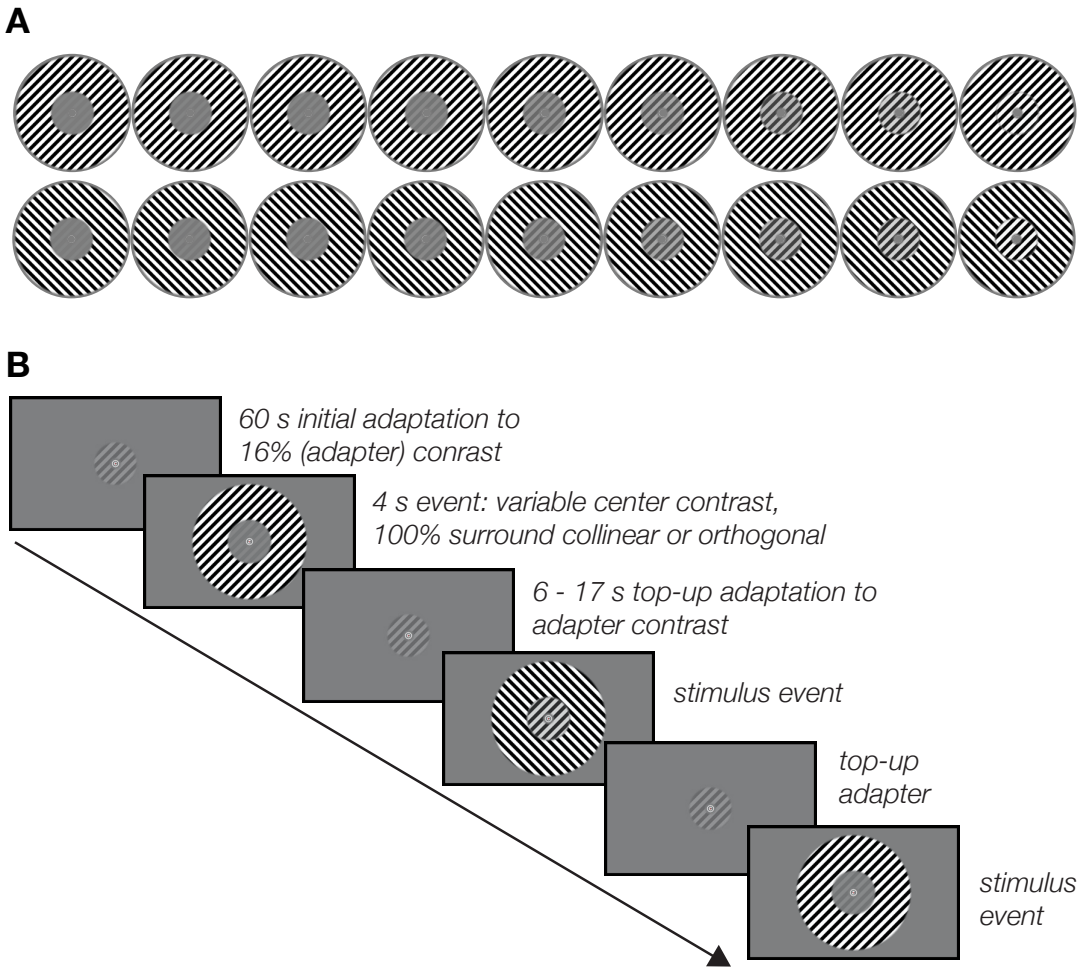

Figure 4.1 Experimental stimuli and paradigm.

(A) Experimental stimuli. Center contrast increases from left to right. Upper row: collinear surround, lower row: orthogonal surround. (B) Three example trials occurring at the start of an fMRI run. Following the 60 s adaptation period, trial order is pseudo-randomized, and inter-trial intervals serve as top-up adapters. In these examples, the center orientation is 45°. Note that spatial frequency was lowered for better visibility.

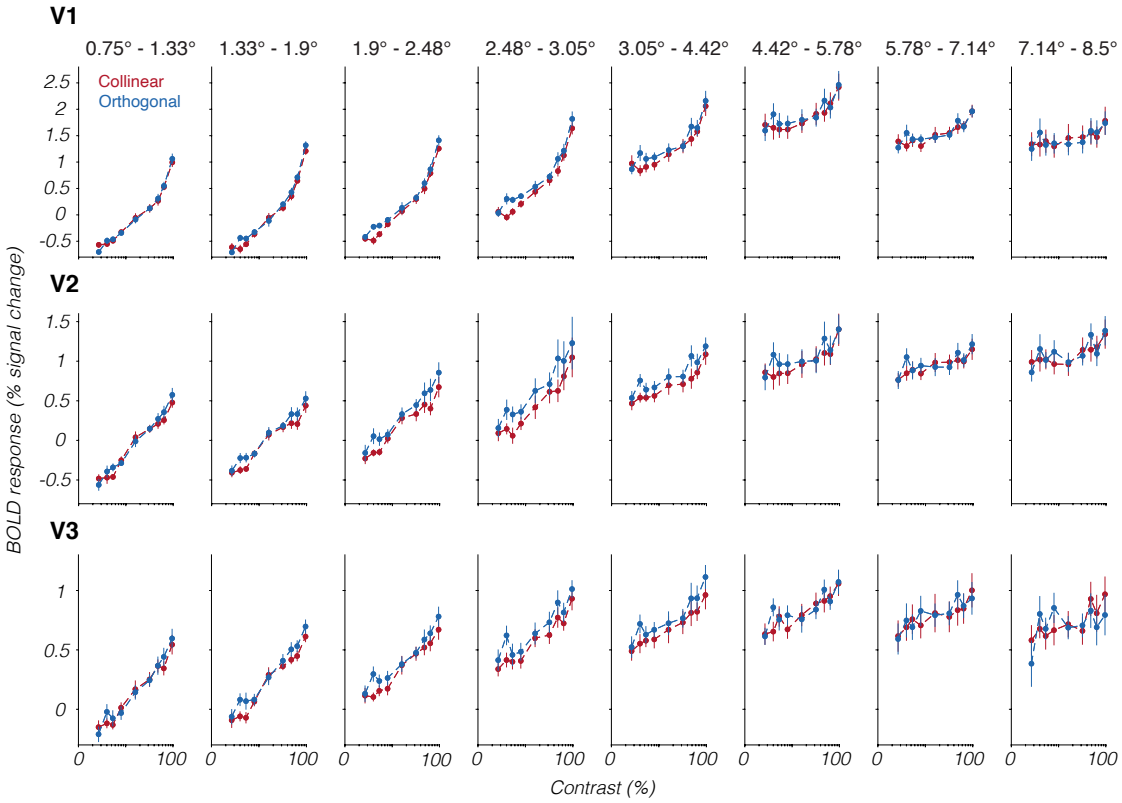

Figure 4.2 Averaged binned contrast responses.

Each row represents results from one visual area; V1: upper row, V2: middle row, V3: bottom row. Left four columns represent the four eccentricity bins into which the center stimulus was divided, right four columns show the four bins of the surround annulus. The bounds of each eccentricity bin are listed above the columns. Center-surround boundary is at 3.05°. The average % signal change across all voxels per observer (n = 10) was calculated first, and the between-observer averages in each condition (red: collinear, blue: orthogonal) are shown in each bin. Error bars represent ± 1 SEM.

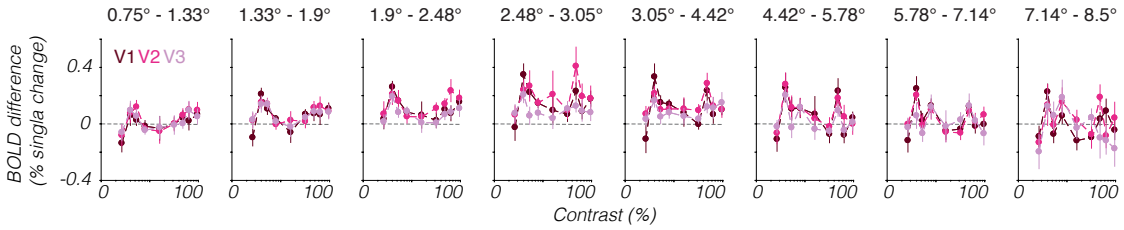

Figure 4.3 Averaged binned orthogonal - collinear BOLD differences.

BOLD % signal change differences between conditions were calculated voxel-wise and then binned into 8 eccentricity bins (4 inner, 4 outer bins; see Figure 4.2). Lines represent the observer-averaged (n = 10) mean difference % signal change values in each eccentricity bin, for each ROI. The largest differences between conditions appear in the center bin closest to boundary (2.48° – 3.05°). Error bars show ± 1 SEM.

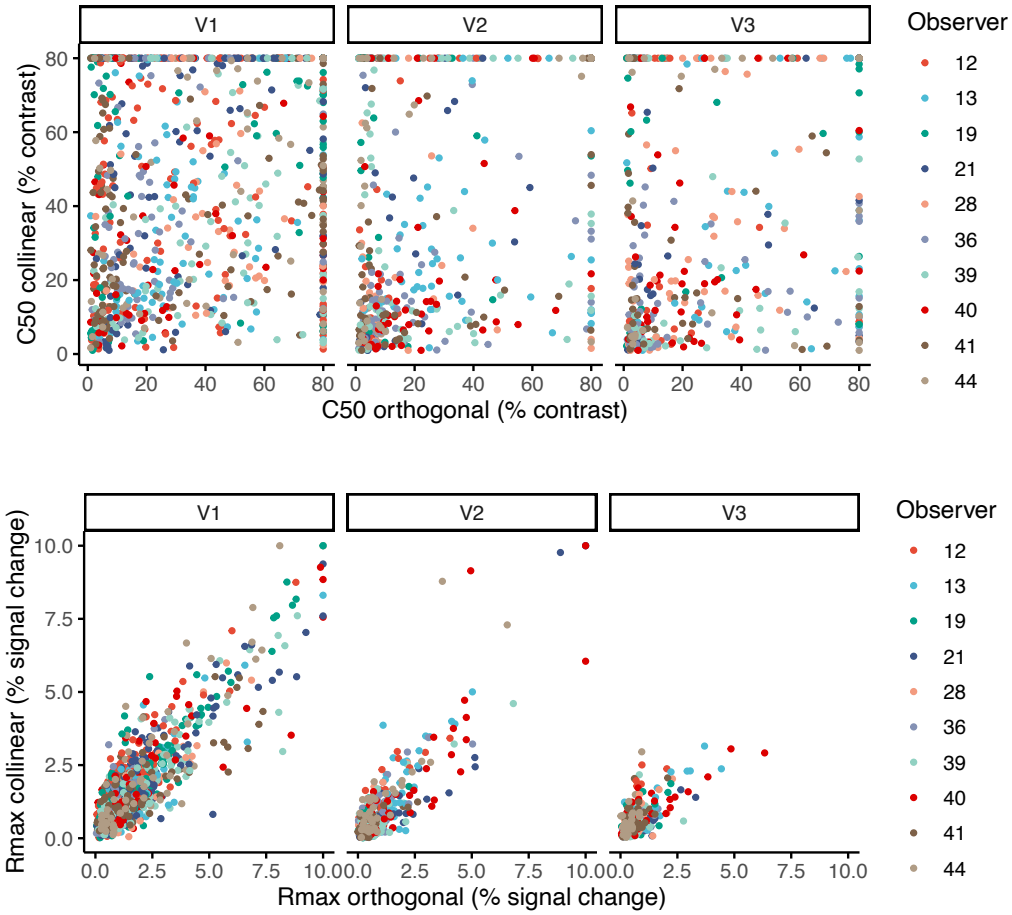

Figure 4.4 Semi-saturation and maximal response scatterplots.

Scatterplots depicting individual voxel's Naka-Rushton parameter values in the collinear surround condition (y-axis) against the orthogonal surround condition (x-axis). Each point represents a single voxel, and data are color-coded by observer (n = 10). Upper panel: C_{50} parameters in collinear vs. orthogonal condition. Bottom panel: R_{max} parameters in collinear vs. orthogonal condition.

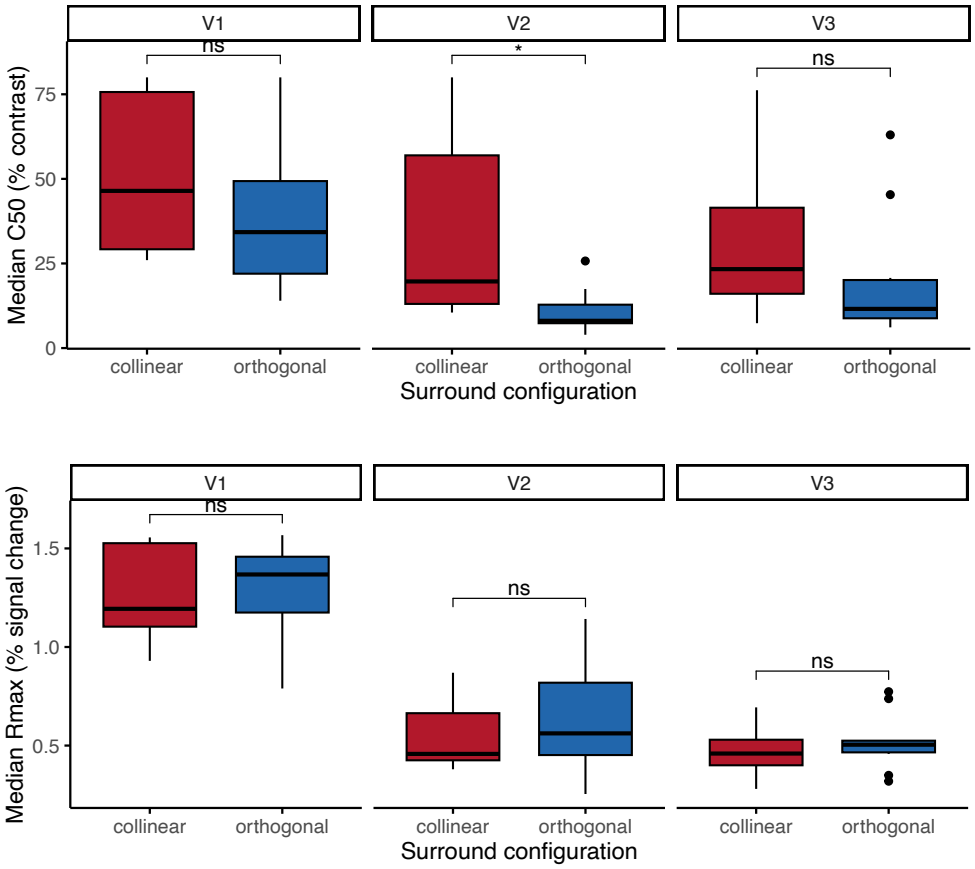

Figure 4.5 Semi-saturation and maximal response statistics.

Boxplots showing median Naka-Rushton estimates across observers (n = 10) compared between surround configurations, within each ROI. Wilcoxon test results are indicated above the plots. n.s. indicates not significant, * indicated $p < 0.05$. Upper row: C_{50} estimates. Bottom row: R_{max} estimates.

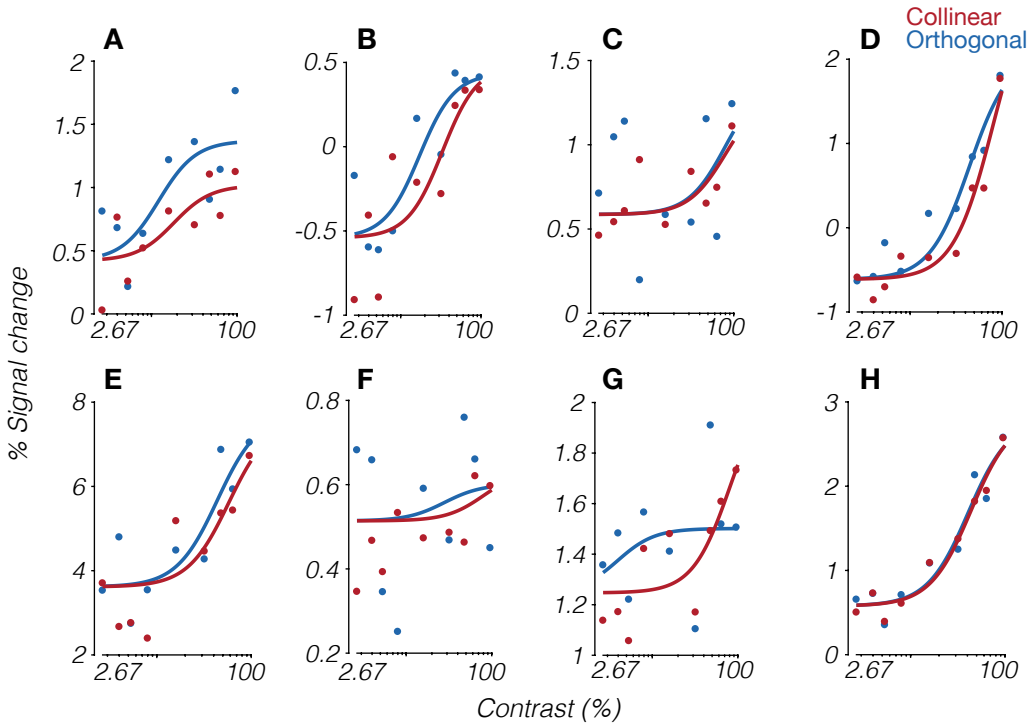

Figure 4.6 Example individual voxel fits.

Data points in each panel show example individual voxel % signal change (PSC) data as a function of center stimulus contrast, while smooth lines show the best-fitting Naka-Rushton in each condition. Data from collinear surround configuration are shown in red, orthogonal in blue. Example voxels were selected to show the variability in CRFs. Voxels A and B show saturating responses, with A showing signs of response gain, and B of contrast gain. Voxel H shows some saturation with overlapping responses between conditions. Voxels C, D, E, and F exhibit non-saturation (with voxel F also showing less responsiveness to contrast and a poorer Naka-Rushton fit), and voxel G shows early saturation in the orthogonal condition but lack of saturation in orthogonal.

CHAPTER FIVE: SUMMARY AND DISCUSSION

Summary of results

The purpose of this dissertation was to investigate the characteristics of orientation-tuned suppression in the human early visual cortex using fMRI. In Chapter 2, we aimed to map the bandwidth of orientation selectivity of tuned normalization. To this end, we measured a full orientation-tuned suppression function with an fMRI blocked design experiment. We used full-field achromatic circular stimuli made up of two sets of spoke components filled with bandpass-filtered white noise whose orientation content could be controlled separately, thus giving us the ability to parametrically manipulate the relative orientations between stimulus components. We measured BOLD responses to seven equally-spaced orientation differences spanning between parallel and orthogonal, and fit a voxel-wise tuning (Gaussian) function to averaged BOLD measurements at each orientation difference, giving us an estimate of the bandwidth of orientation-tuned normalization in the three early visual areas V1, V2, and V3. The results of this first study replicate previous fMRI findings in which collinear stimulus orientations cause stronger BOLD signal suppression than orthogonal (Chen, 2014; Joo et al., 2012; McDonald et al., 2009; Pihlaja et al., 2008; Poltoratski et al., 2017; Schallmo et al., 2016; Williams et al., 2003), but more importantly, they further our knowledge of how suppression strength relates more precisely to orientation differences in human early visual cortex. Suppression strength decreased systematically with increasing orientation differences, consistent with efficient coding of natural scene statistics and suppression of redundant information carried by similar nearby scene locations (Coen-Cagli et al.,

2015). We estimated across 10 observers that the average bandwidth of orientation-tuned suppression is in the range of 20-30° across early visual areas.

In Chapter 3, we addressed the question of whether directing attention towards the visual stimulus changes the properties of the orientation-tuned suppression function measured, most importantly, its bandwidth. We measured BOLD responses to a similar composite stimulus as that in Chapter 2, whose components could have one of six orientation differences between them, under two attention conditions. Participants either held central fixation and performed a fixation task while passively viewing the full-field circular stimuli (unattended condition), while in the attended condition, their task was to hold fixation but monitor the stimulus and report the color of a faint color patch which could sometimes appear anywhere on the stimulus. Fitting a Gaussian tuning function to the deconvolved BOLD responses at each orientation difference in the two attention conditions revealed an overall increase in BOLD response in the attended condition (consistent with previous fMRI attention reports in the literature (Buracas & Boynton, 2007; Itthipuripat et al., 2019; Kastner, 1998; Murray, 2008; Somers et al., 1999)). Importantly, the BOLD response increase was identical across all orientation differences, indicating that attention does not alter the bandwidth of orientation-tuned suppression in early visual cortices. In the previous chapter, we observed that the suppression dependency on orientation difference we observed with BOLD data appears consistent with prior studies mapping natural image regularities; thus, it may be advantageous not to alter the bandwidth of orientation-tuned suppression.

Finally, in Chapter 4, we focused on the effects of orientation-tuned normalization on the early visual contrast response function (CRF), by measuring a central grating CRF to 9 contrast levels under either a collinear or orthogonal high-contrast surround annulus using an event-related design. Voxel-wise deconvolved BOLD responses in each surround condition gave rise to two separate CRF measurements per voxel, which were fit with the Naka-Rushton function, a standard mathematical description of the CRF (Albrecht & Hamilton, 1982). Comparing the two CRF functions revealed suppressive effects, with stronger suppression by collinear surround vs. orthogonal, in agreement with previous findings. Overall, the CRF modulations were most prominent in voxels whose population receptive fields overlapped with the center-surround boundary on the stimulus display, and the statistical summary of results suggests a contrast gain, or a rightward shift of the semi-saturation constant, as opposed to a response gain, or a decrease in responsivity. These results agree with some prior electrophysiological studies exploring CRF modulations under varying surround orientations (DeAngelis et al., 1994; Webb et al., 2005). However, just as electrophysiological investigations show large variability across single cells, we saw substantial CRF modulation variability across voxels, combined with notable lack of saturation in many of the voxels. The results of Chapter 4, while noisy on the voxel-wise level, suggest that the orientation-dependent component of surround suppression in the human early visual areas V1-V3 causes primarily a contrast gain, akin to high contrast adaptation.

Discussion

The experiments contained in this dissertation contribute to the surround modulation fMRI literature in the level of detail provided about the orientation-difference tuning in surround suppression (Chapter 2), the nature of interaction between orientation-tuned surround suppression and visual attention (Chapter 3), and the nature of modulation of the population contrast response function by orientation-tuned surround suppression (Chapter 4). Our findings further demonstrate that suppression in early visual cortex is strongly dependent on the degree of feature differences within the visual display.

Detecting discontinuities in contrast, orientation, and other low-level features in the visual scene is an essential part of scene processing and segmentation, a point of agreement between a wide variety of empirical and computational approaches. Surround modulation is thought to be one of the mechanisms carrying out this function (Sillito et al., 1995). It is widely accepted that neurons in early visual areas are sensitive to contextual modulation in ways which regulate sensitivity and suppress responses to scene locations with uniform texture, while amplifying responses where the RF and the surrounding region differ in contrast, orientation or other low-level features (Levitt & Lund, 1997; Müller et al., 2003). An overarching computational principle is the idea that the visual system tries to reduce redundancy in the neural code, by decreasing correlations between neighboring neurons, especially if they are stimulated with similar textures (Olshausen & Field, 2005; Simoncelli & Olshausen, 2001). One way to achieve this is through adaptation of visual responses to the prevalent natural scene statistics, and the nature and feature sensitivity of contextual modulation suggests this may be the case

(Angelucci et al., 2017; Olshausen & Field, 1997; Vinje & Gallant, 2000). Our results corroborate these ideas, as the bandwidth of orientation-tuned suppression in early visual cortex was relatable to orientation distributions in natural scenes (Sigman et al., 2001) as well as motion direction bandwidth from a previous psychophysical study (Phillips et al., 2021). The lack of bandwidth change with visual attention in Chapter 3 is also compatible with efficient coding principles; if the bandwidth of orientation-tuned suppression is adapted to the biases in co-occurring orientations in natural scenes, it is perhaps not advantageous to alter it.

While it is largely agreed upon that the visual system uses contextual modulation to achieve sparser coding, the neural mechanisms implementing these computations are a matter of much research and debate. However, it is widely assumed that a divisive (as opposed to subtractive) gain control computation takes place, whereby the response of the central RF is modulated based on the activation levels of surrounding neurons (Carandini et al., 1997; Carandini & Heeger, 2012; Cavanaugh et al., 2002a; Fang et al., 2021; Heeger & Zemlianova, 2020). Indeed, a divisive process has been shown in computational studies to decorrelate outputs from V1 neurons modeled on electrophysiological data presented with natural images (Schwartz & Simoncelli, 2001).

The divisive normalization model, first introduced in the early 1990s, succeeds in accounting for a number of nonlinear behaviors in early visual neurons (Carandini et al., 1997; Carandini & Heeger, 2012; Cavanaugh et al., 2002a; Heeger, 1992; Webb et al., 2005), and has also been used successfully to model psychophysical performance (Petrov et al., 2005; Xing & Heeger, 2000, 2001), as well as results obtained with population-

level measurements such as EEG or fMRI (Schallmo et al., 2018; Tsai et al., 2012). While influential and widely used in vision research, a computational study showed that the regular variant of the model performs better with synthetic stimuli as opposed to natural images; however, the model fit was improved by gating the divisive normalization step based on the degree of feature similarity (Coen-Cagli et al., 2015). Chapter 2 results present an informative potential addition to such models to enable more accurate suppression strength modulation based on orientation difference, if the bandwidth were to be incorporated.

The normalization model focuses on the divisive computation which underlies surround modulation, and does not address the anatomical and circuit substrates that carry it out. It is considered highly likely that normalization arises from multiple concurrently operating circuits and mechanism (Carandini & Heeger, 2012; Fang et al., 2021; Heeger & Zemlianova, 2020), but the normalization model has been critiqued for not separating the effects of inhibitory and excitatory connections, the differences between which are an important principle in circuit models based on known anatomical connections between and within visual areas (Ichida et al., 2007; Somers et al., 1998). To this end, electrophysiological and anatomical studies have been critical in mapping different types of suppression within the early visual areas (particularly V1) and their temporal development, and thereby informing the biological constraints of the origins and extent of surround modulation mechanisms. Across studies, an emerging framework focusing on V1 suggests that untuned surround modulation is a product of feedforward connections from the LGN, while horizontal connections, which link preferentially

neurons with similar orientation preferences, and feedback connections from higher areas such as V2/V3, together contribute to the later, feature-tuned components of surround modulation (Angelucci et al., 2017; Angelucci & Bressloff, 2006; Schallmo et al., 2016; Webb et al., 2005). In addition to surround modulation, top-down connections are also thought to carry attentional signals, which interact with the center-surround mechanisms generating surround modulations (Raizada & Grossberg, 2003). Our Chapter 3 results shed some light on the effects of top-down attention on the orientation tuning of surround suppression, namely that the fundamental orientation-tuned suppression bandwidth seems to be preserved under attentional modulation, across early visual areas.

Overall, the majority of electrophysiological and modeling work tends to study V1, and there is less literature focused on surround modulation in later visual areas such as V2 and V3. Area V2 receives its feedforward input from V1 and is thought to be more sensitive to higher-order image structure, such as the orientation differences examined in this dissertation (Freeman et al., 2013). Data from Chapter 2 shows stronger orientation-tuned suppression in area V2, which supports this idea; however, in Chapter 3 we saw no differences between visual areas in orientation-tuned suppression strength. In Chapter 4, differences between collinear and orthogonal surround appear larger in V2 and V3 compared to V1; this could be the result of larger population receptive fields, or the purported increased sensitivity of later areas to higher-order statistics.

Future directions

The stimuli and paradigms used in this study largely succeed in enabling us to characterize the measures of interest, and can be built upon in future experiments. The

investigation of visual cortex responses to changes in orientation and contrast differences is essential in building a roadmap to understanding the broader properties of contextual modulation in visual cortex. Due to the inherent limitations of fMRI study design, in Chapters 2 and 3, we compared BOLD responses to varied orientation differences but held their relative contrast at a constant value. In Chapter 4, we measured a full CRF with 9 contrast levels but contrasting only two orientation difference conditions, collinear vs. orthogonal, as opposed to a finer-grained sequence of orientation difference, which would have given us a more detailed picture of the tuning of orientation-dependent suppression under these conditions. Furthermore, we know that cells in early visual cortex are tuned to multiple features, not only orientation, and respond differently depending not only on contrast and orientation differences, but also stimulus size and spatial frequency. Detailed measurements accounting for richer combination of features would require large numbers of scanning sessions per observer with current fMRI designs. Fast event-related and model-based fMRI approaches (Aghajari et al., 2020; Vinke et al., 2022), which are less time-consuming, may be better suited to fully characterize the landscape of contextual modulation in visual cortex and the interplay of feature sensitivities in contrast, orientation, motion direction, spatial frequency, and more, combined with our ability to measure voxels' population receptive fields. Furthermore, the role of natural scene statistics is generally explored less in the literature, as compared to responses to synthetic visual stimuli and first-order scene statistics; however, there is some compelling evidence that the presence of surrounds that conform to natural scene statistics alters early visual responses and neuron RFs, as compared to

synthetic stimuli (Coen-Cagli et al., 2015; David et al., 2004; Olshausen & Field, 2005). Greater use of natural stimuli combined with normalization models which incorporate feature sensitivity estimates (such as the orientation-tuned suppression bandwidth measured in Chapters 2 and 3) could provide an avenue to better understanding the origins and purpose of early visual cortical responses and their modulation by context.

BIBLIOGRAPHY

Aghajari, S., Vinke, L. N., & Ling, S. (2020). Population spatial frequency tuning in human early visual cortex. Journal of Neurophysiology, 123(2), 773–785. https://doi.org/10.1152/jn.00291.2019

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. Journal of Neurophysiology, 48(1), 217–237. https://doi.org/10.1152/jn.1982.48.1.217

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. NeuroImage, 20(2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

Angelucci, A., Bijanzadeh, M., Nurminen, L., Federer, F., Merlin, S., & Bressloff, P. C. (2017). Circuits and Mechanisms for Surround Modulation in Visual Cortex. Annual Review of Neuroscience, 40(1), 425–451. https://doi.org/10.1146/annurev-neuro-072116-031418

Angelucci, A., & Bressloff, P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. In Progress in Brain Research (Vol. 154, pp. 93–120). Elsevier. https://doi.org/10.1016/S0079-6123(06)54005-1

Atick, J. J. (1992). Could information theory provide an ecological theory of sensory processing? Network: Computation in Neural Systems, 3(2), 213–251. https://doi.org/10.1088/0954-898X_3_2_009

Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time Course and Time-Distance Relationships for Surround Suppression in Macaque V1 Neurons. The Journal of Neuroscience, 23(20), 7690–7701. https://doi.org/10.1523/JNEUROSCI.23-20-07690.2003

Bell, J., & Badcock, D. R. (2008). Luminance and contrast cues are integrated in global shape detection with contours. Vision Research, 48(21), 2336–2344. https://doi.org/10.1016/j.visres.2008.07.015

Benson, N. C., Jamison, K. W., Arcaro, M. J., Vu, A. T., Glasser, M. F., Coalson, T. S., Van Essen, D. C., Yacoub, E., Ugurbil, K., Winawer, J., & Kay, K. (2018). The Human Connectome Project 7 Tesla retinotopy dataset: Description and population receptive field analysis. Journal of Vision, 18(13), 23. https://doi.org/10.1167/18.13.23

Bloem, I. M., & Ling, S. (2019). Normalization governs attentional modulation within human visual cortex. Nature Communications, 10(1), 5660. https://doi.org/10.1038/s41467-019-13597-1

Bloem, I. M., Watanabe, Y. L., Kibbe, M. M., & Ling, S. (2018). Visual Memories Bypass Normalization. Psychological Science, 29(5), 845–856. https://doi.org/10.1177/0956797617747091

Boynton, G. M., Demb, J. B., Glover, G. H., & Heeger, D. J. (1999). Neuronal basis of contrast discrimination. Vision Research, 39(2), 257–269. https://doi.org/10.1016/S0042-6989(98)00113-8

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10(4), 433–436. https://doi.org/10.1163/156856897X00357

Buracas, G. T., & Boynton, G. M. (2007). The Effect of Spatial Attention on Contrast Response Functions in Human Visual Cortex. Journal of Neuroscience, 27(1), 93–97. https://doi.org/10.1523/JNEUROSCI.3162-06.2007

Cannon, M. W., & Fullenkamp, S. C. (1991). Spatial interactions in apparent contrast: Inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. Vision Research, 31(11), 1985–1998. https://doi.org/10.1016/0042-6989(91)90193-9

Cannon, M. W., & Fullenkamp, S. C. (1993). Spatial interactions in apparent contrast: Individual differences in enhancement and suppression effects. Vision Research, 33(12), 1685–1695. https://doi.org/10.1016/0042-6989(93)90034-T

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. Nature Reviews Neuroscience, 13(1), 51–62. https://doi.org/10.1038/nrn3136

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and Normalization in Simple Cells of the Macaque Primary Visual Cortex. The Journal of Neuroscience, 17(21), 8621–8644. https://doi.org/10.1523/JNEUROSCI.17-21-08621.1997

Carrasco, M. (2011). Visual attention: The past 25 years. Vision Research, 51(13), 1484–1525. https://doi.org/10.1016/j.visres.2011.04.012

Carrasco, M., & Barbot, A. (2019). Spatial attention alters visual appearance. Current Opinion in Psychology, 29, 56–64. https://doi.org/10.1016/j.copsyc.2018.10.010

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. Nature Neuroscience, 7(3), 308–313. https://doi.org/10.1038/nn1194

Carrasco, M., Talgar, C. P., & Cameron, E. L. (2001). Characterizing visual performance fields: Effects of transient covert attention, spatial frequency, eccentricity, task and set size. Spatial Vision, 15(1), 61–75. https://doi.org/10.1163/15685680152692015

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and Interaction of Signals From the Receptive Field Center and Surround in Macaque V1 Neurons. Journal of Neurophysiology, 88(5), 2530–2546. https://doi.org/10.1152/jn.00692.2001

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and Spatial Distribution of Signals From the Receptive Field Surround in Macaque V1 Neurons. Journal of Neurophysiology, 88(5), 2547–2556. https://doi.org/10.1152/jn.00693.2001

Chen, C.-C. (2014). Partitioning two components of BOLD activation suppression in flanker effects. Frontiers in Neuroscience, 8. https://doi.org/10.3389/fnins.2014.00149

Coen-Cagli, R., Dayan, P., & Schwartz, O. (2012). Cortical Surround Interactions and Perceptual Salience via Natural Scene Statistics. PLoS Computational Biology, 8(3), e1002405. https://doi.org/10.1371/journal.pcbi.1002405

Coen-Cagli, R., Kohn, A., & Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. Nature Neuroscience, 18(11), 1648–1655. https://doi.org/10.1038/nn.4128

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. Human Brain Mapping, 8(2–3), 109–114. https://doi.org/10.1002/(SICI)1097-0193(1999)8:2/3<109::AID-HBM7>3.0.CO;2-W

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural Stimulus Statistics Alter the Receptive Field Structure of V1 Neurons. The Journal of Neuroscience, 24(31), 6991–7006. https://doi.org/10.1523/JNEUROSCI.1422-04.2004

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. Vision Research, 22(5), 545–559. https://doi.org/10.1016/0042-6989(82)90113-4

DeAngelis, G. C., Freeman, R. D., & Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. Journal of Neurophysiology, 71(1), 347–374. https://doi.org/10.1152/jn.1994.71.1.347

DeAngelis, G. C., Robson, J. G., Ohzawa, I., & Freeman, R. D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex. Journal of Neurophysiology, 68(1), 144–163. https://doi.org/10.1152/jn.1992.68.1.144

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. Annual Review of Neuroscience, 18(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How Does the Brain Solve Visual Object Recognition? Neuron, 73(3), 415–434. https://doi.org/10.1016/j.neuron.2012.01.010

Fang, Z., Bloem, I., Olsson, C., Ma, W. J., & Winawer, J. (2021). Normalization by orientation-tuned surround in human V1-V3. BioRxiv. https://doi.org/10.1101/2021.11.06.467486

Fischl, B. (2012). FreeSurfer. NeuroImage, 62(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Flevaris, A. V., & Murray, S. O. (2015). Attention Determines Contextual Enhancement versus Suppression in Human Primary Visual Cortex. Journal of Neuroscience, 35(35), 12273–12280. https://doi.org/10.1523/JNEUROSCI.1409-15.2015

Foster, J. J., & Ling, S. (2022). Feature-Based Attention Multiplicatively Scales the fMRI-BOLD Contrast-Response Function. The Journal of Neuroscience, 42(36), 6894–6906. https://doi.org/10.1523/JNEUROSCI.0513-22.2022

Freeman, J., Ziemba, C. M., Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (2013). A functional and perceptual signature of the second visual area in primates. Nature Neuroscience, 16(7), 974–981. https://doi.org/10.1038/nn.3402

Gattass, R., Sousa, A., & Gross, C. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. The Journal of Neuroscience, 8(6), 1831–1845. https://doi.org/10.1523/JNEUROSCI.08-06-01831.1988

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. Vision Research, 41(6), 711–724. https://doi.org/10.1016/S0042-6989(00)00277-7

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. NeuroImage, 48(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Gross, C. G., Bender, D. B., & Rocha-Miranda, C. E. (1969). Visual Receptive Fields of Neurons in Inferotemporal Cortex of the Monkey. Science, 166(3910), 1303–1306. https://doi.org/10.1126/science.166.3910.1303

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. Visual Neuroscience, 9(2), 181–197. https://doi.org/10.1017/S0952523800009640

Heeger, D. J., & Zemlianova, K. O. (2020). A recurrent circuit implements normalization, simulating the dynamics of V1 activity. Proceedings of the National Academy of Sciences, 117(36), 22494–22505. https://doi.org/10.1073/pnas.2005417117

Henry, C. A., Joshi, S., Xing, D., Shapley, R. M., & Hawken, M. J. (2013). Functional Characterization of the Extraclassical Receptive Field in Macaque V1: Contrast, Orientation, and Temporal Dynamics. The Journal of Neuroscience, 33(14), 6230–6242. https://doi.org/10.1523/JNEUROSCI.4155-12.2013

Hibbard, P. B., & O'Hare, L. (2015). Uncomfortable images produce non-sparse responses in a model of primary visual cortex. Royal Society Open Science, 2(2), 140535. https://doi.org/10.1098/rsos.140535

Himmelberg, M. M., Winawer, J., & Carrasco, M. (2020). Stimulus-dependent contrast sensitivity asymmetries around the visual field. Journal of Vision, 20(9), 18. https://doi.org/10.1167/jov.20.9.18

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. The Journal of Physiology, 195(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

Ichida, J. M., Schwabe, L., Bressloff, P. C., & Angelucci, A. (2007). Response Facilitation From the "Suppressive" Receptive Field Surround of Macaque V1 Neurons. Journal of Neurophysiology, 98(4), 2168–2181. https://doi.org/10.1152/jn.00298.2007

Itthipuripat, S., Ester, E. F., Deering, S., & Serences, J. T. (2014). Sensory Gain Outperforms Efficient Readout Mechanisms in Predicting Attention-Related Improvements in Behavior. Journal of Neuroscience, 34(40), 13384–13398. https://doi.org/10.1523/JNEUROSCI.2277-14.2014

Itthipuripat, S., Sprague, T. C., & Serences, J. T. (2019). Functional MRI and EEG Index Complementary Attentional Modulations. The Journal of Neuroscience, 39(31), 6162–6179. https://doi.org/10.1523/JNEUROSCI.2519-18.2019

Joo, S. J., Boynton, G. M., & Murray, S. O. (2012). Long-Range, Pattern-Dependent Contextual Effects in Early Human Visual Cortex. Current Biology, 22(9), 781–786. https://doi.org/10.1016/j.cub.2012.02.067

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. Neuron, 15(4), 843–856. https://doi.org/10.1016/0896-6273(95)90175-2

Kastner, S. (1998). Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI. Science, 282(5386), 108–111. https://doi.org/10.1126/science.282.5386.108

Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of Sensory Suppression: Implications for Receptive Field Sizes in the Human Visual Cortex. Journal of Neurophysiology, 86(3), 1398–1411. https://doi.org/10.1152/jn.2001.86.3.1398

Kay, K. N., Winawer, J., Mezer, A., & Wandell, B. A. (2013). Compressive spatial summation in human visual cortex. Journal of Neurophysiology, 110(2), 481–494. https://doi.org/10.1152/jn.00105.2013

Kınıklıoğlu, M., & Boyaci, H. (2022). Increasing the spatial extent of attention strengthens surround suppression. Vision Research, 199, 108074. https://doi.org/10.1016/j.visres.2022.108074

Klímová, M., Bloem, I. M., & Ling, S. (2021). The specificity of orientation-tuned normalization within human early visual cortex. Journal of Neurophysiology, 126(5), 1536–1546. https://doi.org/10.1152/jn.00203.2021

Knierim, J. J., & Van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. Journal of Neurophysiology, 67(4), 961–980. https://doi.org/10.1152/jn.1992.67.4.961

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching Categorical Object Representations in Inferior

Temporal Cortex of Man and Monkey. Neuron, 60(6), 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Kupers, E. R., Carrasco, M., & Winawer, J. (2019). Modeling visual performance differences 'around' the visual field: A computational observer approach. PLOS Computational Biology, 15(5), e1007063. https://doi.org/10.1371/journal.pcbi.1007063

Lee, J., & Maunsell, J. H. R. (2009). A Normalization Model of Attentional Modulation of Single Unit Responses. PLoS ONE, 4(2), e4651. https://doi.org/10.1371/journal.pone.0004651

Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. Nature, 387(6628), 73–76. https://doi.org/10.1038/387073a0

Li, Z. (2002). A saliency map in primary visual cortex. Trends in Cognitive Sciences, 6(1), 9–16. https://doi.org/10.1016/S1364-6613(00)01817-9

Ling, S., & Blake, R. (2009). Suppression During Binocular Rivalry Broadens Orientation Tuning. Psychological Science, 20(11), 1348–1355. https://doi.org/10.1111/j.1467-9280.2009.02446.x

Ling, S., Pratte, M. S., & Tong, F. (2015). Attention alters orientation processing in the human lateral geniculate nucleus. Nature Neuroscience, 18(4), 496–498. https://doi.org/10.1038/nn.3967

Louie, K., Grattan, L. E., & Glimcher, P. W. (2011). Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex. Journal of Neuroscience, 31(29), 10627–10639. https://doi.org/10.1523/JNEUROSCI.1237-11.2011

Maloney, R. T., & Clifford, C. W. G. (2015). Orientation anisotropies in human primary visual cortex depend on contrast. NeuroImage, 119, 129–145. https://doi.org/10.1016/j.neuroimage.2015.06.034

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. Current Biology, 14(9), 744–751. https://doi.org/10.1016/j.cub.2004.04.028

Martínez-Trujillo, J. C., & Treue, S. (2002). Attentional Modulation Strength in Cortical Area MT Depends on Stimulus Contrast. Neuron, 35(2), 365–370. https://doi.org/10.1016/S0896-6273(02)00778-X

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. The Journal of Neuroscience, 19(1), 431–441. https://doi.org/10.1523/JNEUROSCI.19-01-00431.1999

McDonald, J. S., Seymour, K. J., Schira, M. M., Spehar, B., & Clifford, C. W. G. (2009). Orientation-specific contextual modulation of the fMRI BOLD response to luminance and chromatic gratings in human visual cortex. Vision Research, 49(11), 1397–1405. https://doi.org/10.1016/j.visres.2008.12.014

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Uğurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magnetic Resonance in Medicine, 63(5), 1144–1153. https://doi.org/10.1002/mrm.22361

Moran, J., & Desimone, R. (1985). Selective Attention Gates Visual Processing in the Extrastriate Cortex. Science, 229(4715), 782–784. https://doi.org/10.1126/science.4023713

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. Journal of Neurophysiology, 70(3), 909–919. https://doi.org/10.1152/jn.1993.70.3.909

Müller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (2003). Local Signals From Beyond the Receptive Fields of Striate Cortical Neurons. Journal of Neurophysiology, 90(2), 822–831. https://doi.org/10.1152/jn.00005.2003

Murray, S. O. (2008). The effects of spatial attention in early human visual cortex are stimulus independent. Journal of Vision, 8(10), 2–2. https://doi.org/10.1167/8.10.2

Naka, K. I., & Rushton, W. A. H. (1966). S-potentials from luminosity units in the retina of fish (Cyprinidae). The Journal of Physiology, 185(3), 587–599. https://doi.org/10.1113/jphysiol.1966.sp008003

Ni, A. M., & Maunsell, J. H. R. (2017). Spatially tuned normalization explains attention modulation variance within neurons. Journal of Neurophysiology, 118(3), 1903–1913. https://doi.org/10.1152/jn.00218.2017

Ni, A. M., Ray, S., & Maunsell, J. H. R. (2012). Tuned Normalization Explains the Size of Attention Modulations. Neuron, 73(4), 803–813. https://doi.org/10.1016/j.neuron.2012.01.006

Ohzawa, I., Sclar, G., & Freeman, R. D. (1985). Contrast gain control in the cat's visual system. Journal of Neurophysiology, 54(3), 651–667. https://doi.org/10.1152/jn.1985.54.3.651

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1? Vision Research, 37(23), 3311–3325. https://doi.org/10.1016/S0042-6989(97)00169-7

Olshausen, B. A., & Field, D. J. (2005). How Close Are We to Understanding V1? Neural Computation, 17(8), 1665–1699. https://doi.org/10.1162/0899766054026639

Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. Nature Neuroscience, 5(12), 1332–1338. https://doi.org/10.1038/972

Petrov, Y., Carandini, M., & McKee, S. (2005). Two Distinct Mechanisms of Suppression in Human Vision. Journal of Neuroscience, 25(38), 8704–8707. https://doi.org/10.1523/JNEUROSCI.2871-05.2005

Phillips, D. J., McDougall, T. J., Dickinson, J. E., & Badcock, D. R. (2021). Motion direction tuning in centre-surround suppression of contrast. Vision Research, 179, 85–93. https://doi.org/10.1016/j.visres.2020.11.001

Pihlaja, M., Henriksson, L., James, A. C., & Vanni, S. (2008). Quantitative multifocal fMRI shows active suppression in human V1. Human Brain Mapping, 29(9), 1001–1014. https://doi.org/10.1002/hbm.20442

Pokorny, V. J., Schallmo, M.-P., Sponheim, S. R., & Olman, C. A. (2023). Weakened untuned gain control is associated with schizophrenia while atypical orientation-

tuned suppression depends on visual acuity. Journal of Vision, 23(2), 2. https://doi.org/10.1167/jov.23.2.2

Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., & Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. Nature, 391(6667), 580–584. https://doi.org/10.1038/35372

Polat, U., & Norcia, A. M. (1996). Neurophysiological Evidence for Contrast Dependent Long-range Facilitation and Suppression in the Human Visual Cortex. Vision Research, 36(14), 2099–2109. https://doi.org/10.1016/0042-6989(95)00281-2

Poltoratski, S., Ling, S., McCormack, D., & Tong, F. (2017). Characterizing the effects of feature salience and top-down attention in the early visual system. Journal of Neurophysiology, 118(1), 564–573. https://doi.org/10.1152/jn.00924.2016

Poltoratski, S., Maier, A., Newton, A. T., & Tong, F. (2019). Figure-Ground Modulation in the Human Lateral Geniculate Nucleus Is Distinguishable from Top-Down Attention. Current Biology, 29(12), 2051-2057.e3. https://doi.org/10.1016/j.cub.2019.04.068

Porciatti, V., Bonanni, P., Fiorentini, A., & Guerrini, R. (2000). Lack of cortical contrast gain control in human photosensitive epilepsy. Nature Neuroscience, 3(3), 259–263. https://doi.org/10.1038/72972

Rabinowitz, N. C., Willmore, B. D. B., Schnupp, J. W. H., & King, A. J. (2011). Contrast Gain Control in Auditory Cortex. Neuron, 70(6), 1178–1191. https://doi.org/10.1016/j.neuron.2011.04.030

Raizada, R. D. S., & Grossberg, S. (2003). Towards a Theory of the Laminar Architecture of Cerebral Cortex: Computational Clues from the Visual System. Cerebral Cortex, 13(1), 100–113. https://doi.org/10.1093/cercor/13.1.100

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. NeuroImage, 53(4), 1181–1196. https://doi.org/10.1016/j.neuroimage.2010.07.020

Reynolds, J. H., & Chelazzi, L. (2004). ATTENTIONAL MODULATION OF VISUAL PROCESSING. Annual Review of Neuroscience, 27(1), 611–647. https://doi.org/10.1146/annurev.neuro.26.041002.131039

Reynolds, J. H., & Heeger, D. J. (2009). The Normalization Model of Attention. Neuron, 61(2), 168–185. https://doi.org/10.1016/j.neuron.2009.01.002

Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention Increases Sensitivity of V4 Neurons. Neuron, 26(3), 703–714. https://doi.org/10.1016/S0896-6273(00)81206-4

Sceniak, M. P., Hawken, M. J., & Shapley, R. (2001). Visual Spatial Characterization of Macaque V1 Neurons. Journal of Neurophysiology, 85(5), 1873–1887. https://doi.org/10.1152/jn.2001.85.5.1873

Schallmo, M.-P., Grant, A. N., Burton, P. C., & Olman, C. A. (2016). The effects of orientation and attention during surround suppression of small image features: A 7 Tesla fMRI study. Journal of Vision, 16(10), 19. https://doi.org/10.1167/16.10.19

Schallmo, M.-P., Kale, A. M., Millin, R., Flevaris, A. V., Brkanac, Z., Edden, R. A., Bernier, R. A., & Murray, S. O. (2018). Suppression and facilitation of human neural responses. ELife, 7, e30334. https://doi.org/10.7554/eLife.30334

Schallmo, M.-P., Kale, A. M., & Murray, S. O. (2019). The time course of different surround suppression mechanisms. Journal of Vision, 19(4), 12. https://doi.org/10.1167/19.4.12

Schallmo, M.-P., & Murray, S. O. (2016). Identifying separate components of surround suppression. Journal of Vision, 16(1), 2. https://doi.org/10.1167/16.1.2

Schallmo, M.-P., Sponheim, S. R., & Olman, C. A. (2013). Abnormal Contextual Modulation of Visual Contour Detection in Patients with Schizophrenia. PLoS ONE, 8(6), e68090. https://doi.org/10.1371/journal.pone.0068090

Schwabe, L., Obermayer, K., Angelucci, A., & Bressloff, P. C. (2006). The Role of Feedback in Shaping the Extra-Classical Receptive Field of Cortical Neurons: A Recurrent Network Model. The Journal of Neuroscience, 26(36), 9117–9129. https://doi.org/10.1523/JNEUROSCI.1253-06.2006

Schwartz, O., Sejnowski, T. J., & Dayan, P. (2009). Perceptual organization in the tilt illusion. Journal of Vision, 9(4), 19–19. https://doi.org/10.1167/9.4.19

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. Nature Neuroscience, 4(8), 819–825. https://doi.org/10.1038/90526

Self, M. W., Lorteije, J. A. M., Vangeneugden, J., van Beest, E. H., Grigore, M. E., Levelt, C. N., Heimel, J. A., & Roelfsema, P. R. (2014). Orientation-Tuned

Surround Suppression in Mouse Visual Cortex. Journal of Neuroscience, 34(28), 9290–9304. https://doi.org/10.1523/JNEUROSCI.5051-13.2014

Sengpiel, F., Baddeley, R. J., Freeman, T. C. B., Harrad, R., & Blakemore, C. (1998). Different mechanisms underlie three inhibitory phenomena in cat area 17. Vision Research, 38(14), 2067–2080. https://doi.org/10.1016/S0042-6989(97)00413-6

Shushruth, S., Ichida, J. M., Levitt, J. B., & Angelucci, A. (2009). Comparison of Spatial Summation Properties of Neurons in Macaque V1 and V2. Journal of Neurophysiology, 102(4), 2069–2083. https://doi.org/10.1152/jn.00512.2009

Shushruth, S., Mangapathy, P., Ichida, J. M., Bressloff, P. C., Schwabe, L., & Angelucci, A. (2012). Strong Recurrent Networks Compute the Orientation Tuning of Surround Modulation in the Primate Primary Visual Cortex. Journal of Neuroscience, 32(1), 308–321. https://doi.org/10.1523/JNEUROSCI.3789-11.2012

Shushruth, S., Nurminen, L., Bijanzadeh, M., Ichida, J. M., Vanni, S., & Angelucci, A. (2013). Different Orientation Tuning of Near- and Far-Surround Suppression in Macaque Primary Visual Cortex Mirrors Their Tuning in Human Perception. Journal of Neuroscience, 33(1), 106–119. https://doi.org/10.1523/JNEUROSCI.2518-12.2013

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. Proceedings of the National Academy of Sciences, 98(4), 1935–1940. https://doi.org/10.1073/pnas.98.4.1935

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davls, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. Nature, 378(6556), 492–496. https://doi.org/10.1038/378492a0

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural Image Statistics and Neural Representation. Annual Review of Neuroscience, 24(1), 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, 23 Suppl 1, S208-219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. Vision Research, 33(18), 2671–2683. https://doi.org/10.1016/0042-6989(93)90227-N

Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. H. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. Proceedings of the National Academy of Sciences, 96(4), 1663–1668. https://doi.org/10.1073/pnas.96.4.1663

Somers, D. C., Todorov, E. V., Siapas, A. G., Toth, L. J., Kim, D.-S., & Sur, M. (1998). A local circuit approach to understanding integration of long-range inputs in

primary visual cortex. Cerebral Cortex, 8(3), 204–217. https://doi.org/10.1093/cercor/8.3.204

Tolhurst, D. J., & Heeger, D. J. (1997). Comparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex. Visual Neuroscience, 14(2), 293–309. https://doi.org/10.1017/S0952523800011433

Tootell, R. B. H., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., & Dale, A. M. (1998). Functional analysis of primary visual cortex (V1) in humans. Proceedings of the National Academy of Sciences, 95(3), 811–817. https://doi.org/10.1073/pnas.95.3.811

Toth, L. J., Rao, S. C., Kim, D. S., Somers, D., & Sur, M. (1996). Subthreshold facilitation and suppression in primary visual cortex revealed by intrinsic signal imaging. Proceedings of the National Academy of Sciences, 93(18), 9869–9874. https://doi.org/10.1073/pnas.93.18.9869

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Treue, S. (2001). Neural correlates of attention in primate visual cortex. Trends in Neurosciences, 24(5), 295–300. https://doi.org/10.1016/S0166-2236(00)01814-2

Treue, S., & Trujillo, J. C. M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature, 399(6736), 575–579. https://doi.org/10.1038/21176

Trott, A. R., & Born, R. T. (2015). Input-Gain Control Produces Feature-Specific Surround Suppression. Journal of Neuroscience, 35(12), 4973–4982. https://doi.org/10.1523/JNEUROSCI.4000-14.2015

Tsai, J. J., Wade, A. R., & Norcia, A. M. (2012). Dynamics of Normalization Underlying Masking in Human Visual Cortex. Journal of Neuroscience, 32(8), 2783–2789. https://doi.org/10.1523/JNEUROSCI.4485-11.2012

Ungerleider, L. G., Desimone, R., Galkin, T. W., & Mishkin, M. (1984). Subcortical projections of area MT in the macaque. The Journal of Comparative Neurology, 223(3), 368–386. https://doi.org/10.1002/cne.902230304

Vinje, W. E., & Gallant, J. L. (2000). Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. Science, 287(5456), 1273–1276. https://doi.org/10.1126/science.287.5456.1273

Vinke, L. N., Bloem, I. M., & Ling, S. (2022). Saturating Nonlinearities of Contrast Response in Human Visual Cortex. The Journal of Neuroscience, 42(7), 1292–1302. https://doi.org/10.1523/JNEUROSCI.0106-21.2021

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. Perception & Psychophysics, 33(2), 113–120. https://doi.org/10.3758/BF03202828

Webb, B. S., Dhruv, N. T., Solomon, S., Tailby, C., & Lennie, P. (2005). Early and Late Mechanisms of Surround Suppression in Striate Cortex of Macaque. Journal of Neuroscience, 25(50), 11666–11675. https://doi.org/10.1523/JNEUROSCI.3414-05.2005

Webb, B. S., Tinsley, C. J., Barraclough, N. E., Parker, A., & Derrington, A. M. (2003). Gain control from beyond the classical receptive field in primate primary visual cortex. Visual Neuroscience, 20(3), 221–230. https://doi.org/10.1017/S0952523803203011

Williams, A. L., Singh, K. D., & Smith, A. T. (2003). Surround Modulation Measured With Functional MRI in the Human Visual Cortex. J Neurophysiol, 89, 9. https://doi.org/10.1152/jn.00048.2002

Williford, T., & Maunsell, J. H. R. (2006). Effects of Spatial Attention on Contrast Response Functions in Macaque Area V4. Journal of Neurophysiology, 96(1), 40–54. https://doi.org/10.1152/jn.01207.2005

Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. Vision Research, 40(22), 3065–3072. https://doi.org/10.1016/S0042-6989(00)00152-8

Xing, J., & Heeger, D. J. (2001). Measurement and modeling of center-surround suppression and enhancement. Vision Research, 41(5), 571–583. https://doi.org/10.1016/S0042-6989(00)00270-4

Xu, J., Moeller, S., Auerbach, E. J., Strupp, J., Smith, S. M., Feinberg, D. A., Yacoub, E., & Uğurbil, K. (2013). Evaluation of slice accelerations using multiband echo planar imaging at 3T. NeuroImage, 83, 991–1001. https://doi.org/10.1016/j.neuroimage.2013.07.055

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. Journal of Vision, 10(10), 16–16. https://doi.org/10.1167/10.10.16

Zenger-Landolt, B., & Heeger, D. J. (2003). Response Suppression in V1 Agrees with Psychophysics of Surround Masking. The Journal of Neuroscience, 23(17), 6884–6893. https://doi.org/10.1523/JNEUROSCI.23-17-06884.2003

CURRICULUM VITAE