# Neural encoding of object properties and task-dependent changes in primate visual area V4

Dina V. Popovkina

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Anitha Pasupathy, Chair

Wyeth Bair

Gregory Horwitz

Program Authorized to Offer Degree:

Neuroscience

University of Washington

## Abstract

Neural encoding of object properties and task-dependent changes in primate visual area V4

Dina V. Popovkina

Chair of the Supervisory Committee:
Associate Professor Anitha Pasupathy
Department of Biological Structure

Primates use vision to understand and interact with the world around them. Several interconnected visual areas are responsible for processing information about visual objects, ultimately enabling object and scene recognition and perception. Visual area V4 is an intermediate stage in the pathway of object processing, and V4 neurons are sensitive to multiple aspects of objects such as their form, texture, and color.

The first goal of this dissertation was to investigate how responses of neurons in visual area V4 reflect information about object boundaries and interiors. Unlike prominent computational models of object recognition which rely on boundaries, the majority of neurons in area V4 displayed a modulation of responses by the presence or absence of an object interior. I developed two computational model modifications that successfully incorporated both of the response characteristics that I observed

experimentally: sensitivity to boundaries and sensitivity to the presence of an object interior.

The second goal of this dissertation was to examine whether V4 response selectivity for shape and color of visual objects depended on the task which the animal was performing. I compared responses during a task where the animal was discriminating object shape to a task where the animal was passively viewing the same objects. I found that the majority of neurons displayed not only a change in the magnitude of responses between the two tasks, but that response changes were additionally dependent on object shape and color.

Together, these studies contribute to our understanding of how neurons encode visual information, and how this encoding is changed during behavior.

**TABLE OF CONTENTS**

## INDEX OF FIGURES

## INDEX OF EQUATIONS

**ACKNOWLEDGMENTS**

First and foremost, I want to thank the two macaque monkeys, Odin and Puck, whose hard work made this dissertation possible. I'm so grateful to have been able to work with these amazing animals, who have allowed me to glimpse some of the inner workings of their brain, and who were patient with me as I learned to speak in their terms.

I am thankful that my advisor, Anitha Pasupathy, accepted me into her lab although I knew little about physiology, and less about vision. She helped me learn how to think about and carry out vision research from the ground up. She has been an incredible mentor who consistently pushed me to do more and do better, who helped me develop confidence in my abilities while holding on to a healthy dose of skepticism that could take the quality of my work to a higher level. I am inspired by her honesty and conviction, her passion for our science and our questions, and I am deeply indebted to her for advice and support over the years.

I also want to thank members of my dissertation committee, who provided diverse perspectives that reflect the interdisciplinary nature of this work and our field: Wyeth Bair, for the hours he spent helping me get a handle on the modeling aspects of my work; Greg Horwitz and Geoff Boynton, for helping me expand my horizons in vision; Jay Neitz, for inspiring me to think about color in holistic ways; and Eric Shea-Brown, for providing a valuable computational perspective and being a constant source of positivity. Their support, combined with rigorous and thoughtful suggestions, always pushed my projects in productive directions.

My graduate experience was enriched by the members of the Pasupathy lab. I thank Yasmine El-Shamayleh for being an integral part of my training. She has invested a lot of time and effort into my growth, and I feel so lucky that I can call this talented, dedicated person both my mentor and my friend. I'm also happy to have had a solid brunch/dinner buddy over the years, and for the many meals we have spent working on goals/mini goals and embracing our queendom.

I am deeply indebted to Amber Fyall for coaching me past many day-to-day challenges, both in work and in life. I'm grateful for both her monkey whispering skills and her ability to find the right cat photo at the right time. Brittany Bushnell and Yoshito Kosai have both contributed to the crucial first steps of my projects, especially the animal training and task/analysis coding necessary to hit the ground running. I also thank the rest of the lab who have shared in this experience and provided important thoughts and feedback: Tomoyuki Namima, Taekjun Kim, Polina Zamarashkina, Tim Oleskiw, and Amy Nowack; and members of the Bair and Horwitz labs: Pamela Baker, Dean Pospisil, Patrick Weller, and Abhishek De.

I am delighted that I made the choice to come to UW, and that my initial impressions of the happy, engaged students in the Neuroscience program were a reliable predictor of the wonderful peers I have been able to meet. Some of them, like Philia Gau and Alison Mehravari, have become true best friends. I am so grateful for their support through the good times and the bad, for a warm hug or a slice of pie when it was most needed, and for listening to my billions of stories. I am thankful to Leah Bakst and Seth Koenig, who have helped me process the feelings related to the mental, physical, and emotional challenges of primate work, which they know firsthand. I also thank Florie

D’Orazi, Seoan Liu, Sarah Pickett, Clare Gamlin, and Andrea McQuate for their willingness to spend time openly thinking about our path as scientists in between experiments or analysis/writing marathons in the many libraries and conference rooms of the UW campus. I am also thankful to have overlapped in my training with several supportive and inspiring students from older cohorts: Charlie Hass, Bill Wood, Braulio Peguero, and Hiro Watari; and I am thankful to the Perkel and Neitz labs for really great rotation project experiences.

I want to thank the Neuroscience Program staff (Lucia Wisdom, Ann Wilkinson, and Margie Trenary) for helping things run smoothly on the administrative side so I could focus on my science. My work was also made possible by the support of several funding sources, including the NSF Graduate Research Fellowship Program and NIH training grants at UW (Computational Neuroscience and Vision).

I’m thankful for other support I received from outside of my graduate program and even outside UW. Joris Vincent and Kit Moreland helped me get a better understanding of how my non-human primate work relates to human primates. John Palmer provided many additional perspectives on my work, and I am especially grateful for his “big-picture” feedback and the opportunities to improve my communication skills. I am grateful for experience volunteering with the Neuro Outreach program at UW and the Science Communication Fellowship at the Pacific Science Center, which helped me learn to transform complicated concepts into engaging and accessible ways to reach non-experts. I am grateful to have helped a few girls have an awesome time learning about science, and these experiences remind me why I love what I do.

I would not have made it to graduate school at all were it not for a few really special people. First, my mom and dad, who have supported me unconditionally in all of my (sometimes crazy) aspirations; and my siblings, uncles/aunts, cousins, and grandparents. If I appear to have grown into a mostly functional and well-rounded adult, it is completely thanks to their guidance. I owe my scientific inclinations especially to one of my grandmothers, E. M. Popovkina (published under Lipovetskaia), who quietly fed my early curiosity by setting me up with a kids' book about molecular biology ("Journey to the land of MOB", 1967). I am also grateful to my cat Venus, in large part for making sure I get enough rest (she doesn't let me sleep on the couch).

I owe many of my successes to the Meyerhoff Scholarship Program at UMBC, my *alma mater*. I found a home and a giant new family as an M, thanks to my amazing peers (such as Stephanie Battle, who has shared both the MD-WA and the PhD journeys with me) and staff (especially LaMont Toliver, Keith Harmon, and Mitsue Wiggs). Their enthusiastic support led me to spend 3 years in the lab of Scott Thompson at UMB SoM, who taught me to love the brain and guided me to the path of learning how to think about it. I also wanted to thank Scott Dixon, my mentor during a summer I spent away from the Thompson lab, who helped me see that good science is limited only by your imagination.

I am truly fortunate to be surrounded by many other people who accept and support me, friends who are more like family. These include Anna Retherford, my best friend of many, many years and one of the few people with whom I can hold a multi-hour phone conversation; Kirsten and Lucas Taylor, whose door was always open and who graciously kept me sane and well-fed; Jasper Schneider, whose sense of humor matches mine with scary levels of overlap; Steven Stone and Jason Sellers, who have been great enablers

of my extracurricular interests, like watching a ball be kicked around for 90 minutes. I am in the company of some really awesome science friends, including Hayley Mattison, Casey Daniels, and Madineh Sarvestani. Last but not least, I am really thankful to dear friends in the Seattle Argentine tango community, who have taught me a lot about living my best life.

## DEDICATION

*To my family,*

*and my greatest sources of inspiration:*

*people who create beautiful things*

## Chapter 1. Introduction and literature overview

### *1.1 Vision in primates*

Visual information is a critical part of a primate's sensory experience in the world. Primates rely heavily on their sense of vision to survive in complex natural environments as they navigate to avoid danger and obstacles, forage for food, and socialize with others. To enable the multitude of decisions that underlie these diverse behaviors, the primate brain must process sensory input quickly and reliably. Given the dynamic nature of the world and rapidly changing behavioral goals, the functional demands present a significant computational challenge for the visual system. A large portion of the primate brain is devoted to overcoming this hurdle: over half of the primate neocortex is devoted to vision or associated processing (Felleman and Van Essen, 1991).

Below, I will first provide a broad overview of how visual information is organized in the visual system; then, I will highlight how different aspects of objects, such as form, are represented in neuronal responses; and finally, I will focus on the intermediate stages of object processing, which lie at a crossroads between perception and action.

### *1.2 Organization of the primate visual system*

Information processing in the visual system is arranged in stages. Photons entering the eye are first captured by retinal receptors; once the retinal network transforms light into electrical impulses, this information is relayed via the lateral geniculate nucleus of the thalamus (LGN) to the cortex, beginning with the primary visual area (V1). Nearly a century of work (Brodmann 1909; Felleman and Van Essen 1991)

has revealed a constellation of functionally organized and interconnected areas that are ultimately responsible for our perception of the visual world. Processing in visual brain areas can be broadly captured by a hierarchical structure (**Fig. 1.1**).

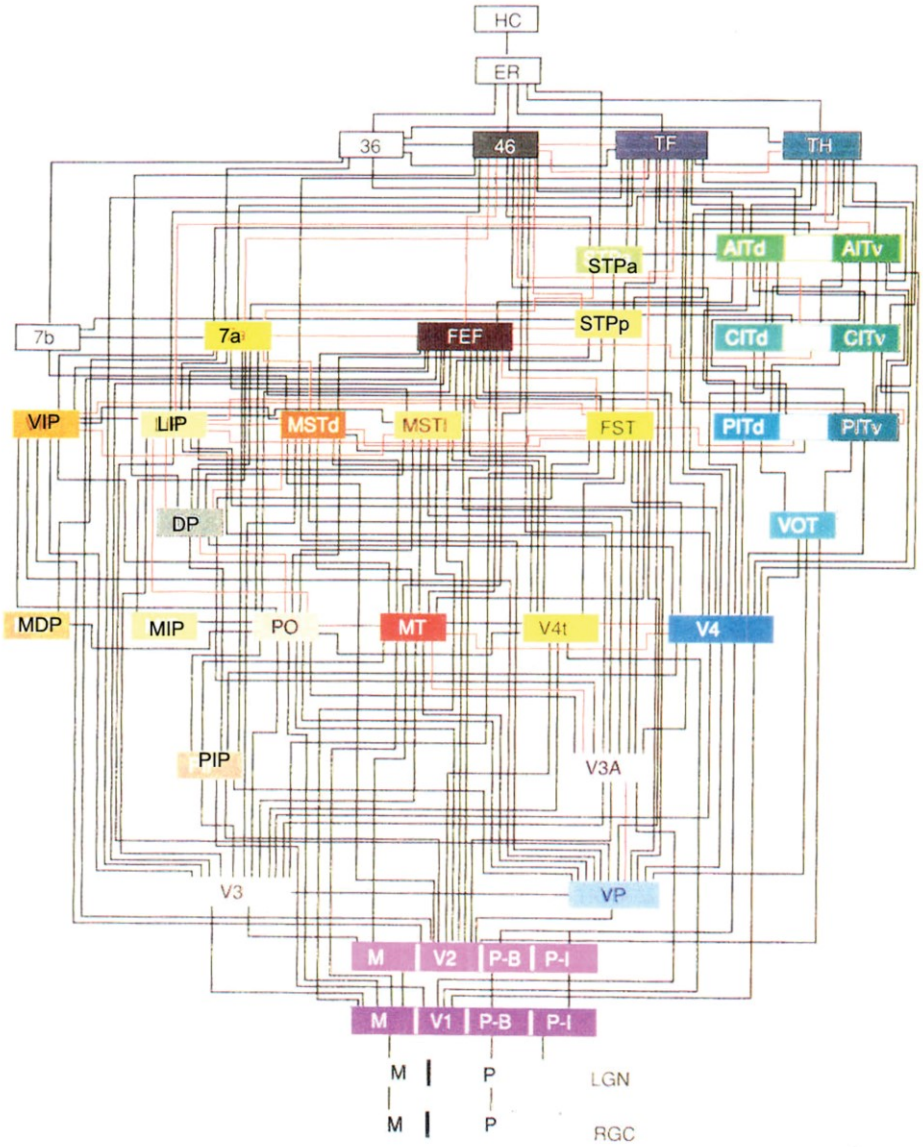


**Figure 1.1.** *Organization of the visual system*.

Reproduced with permission from Felleman and Van Essen (1991). Visual areas are denoted with boxes and connections between areas with lines. This representation is extended to begin with retinal ganglion cells of the retina (RGC) and end with the nonvisual hippocampal complex (HC). This hierarchy is arranged according to the laminarity and anterograde/retrograde patterning of connections; horizontal groups show areas that are at similar levels of processing.

Within this structure, neurons at each successive stage represent a larger area of visual space (spatial receptive field, RF; Hartline, 1938; Perkel *et al.*, 1986; Gattass *et al.*, 1988) and are sensitive to more complex aspects of visual stimuli. For example, neurons may be sensitive to the direction of motion of a bar of light in early visual areas, and to complex optic flow patterns in later areas; or, to the orientation of a bar of light in early areas, and entire objects such as faces in later areas (Van Essen and Gallant, 1994; Kobatake and Tanaka, 1994). Additionally, information about visual object motion appears to be largely segregated from information about object appearance and identity (Goodale and Milner, 1992); although these streams of information are not wholly separated (Goodale and Westwood, 2004; Cloutman, 2013). The subset of areas where neural responses specifically reflect object appearance project along the ventral surface of the primate brain, towards inferotemporal cortex. Collectively, these areas are termed the ventral stream, or pathway, of visual processing (**Fig. 1.2**; Kravitz *et al.*, 2013).

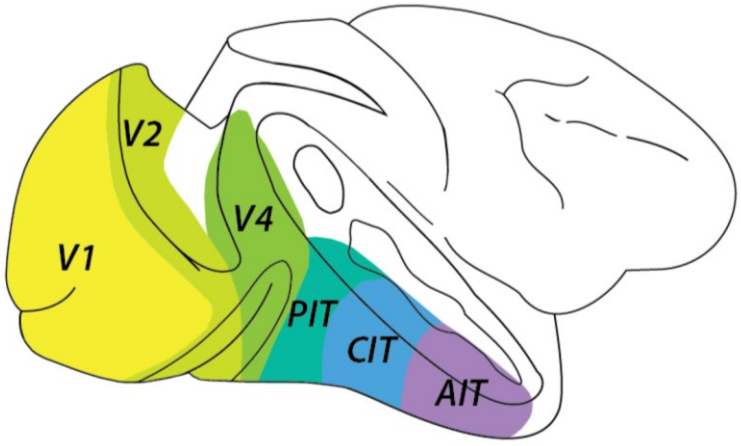


**Figure 1.2.** *Ventral pathway of visual processing*.

Adapted with permission from Parker (2007). In this lateral view of the macaque cerebral cortex, the colored areas are part of the ventral pathway which underlies object perception and recognition. V1: primary visual cortex; PIT: posterior inferotemporal cortex; CIT: central inferotemporal cortex; AIT: anterior inferotemporal cortex.

*1.3 Ventral pathway representation of object form*

As described above, the hierarchical structure of the visual system is reflected in the increasingly complex neural sensitivity to object features as information passes through successive stages. This trend holds true for representation of form in the ventral pathway. For example, neurons in area V1 are selective for orientation and spatial frequency (SF) of visual stimuli (Hubel and Wiesel, 1959, 1968; Movshon *et al.*, 1978a, b; Albrecht *et al.*, 1980), meaning neurons display larger spiking responses to certain orientations and SFs, and not others. At the next stage, area V2, many neurons respond selectively to conjunctions of oriented bars or the angles of their apertures (Ito and Komatsu, 2004). In later stages such as area V4 and inferotemporal cortex (IT), neurons become sensitive to and selective for even more complex forms, such as non-Cartesian gratings (Gallant *et al.*, 1996), partial object boundaries (Pasupathy and Connor, 2001), and faces and hands (Kobatake and Tanaka, 1994).

*1.4 Other object properties represented by neural responses in the ventral pathway*

Neurons at successive stages in the ventral stream represent different visual aspects of objects relating to their surface properties, such as color and texture. For example, the population responses of neurons in area V4 appear to be selective for both stimulus hue and luminance contrast, whereas neurons in anterior portions of IT cortex tend to be invariant to luminance contrast (Namima *et al.*, 2014). This observation suggests that in later stages of the ventral stream, the neural representation of stimulus color becomes more aligned with the behavioral phenomenon of color constancy (reliable recognition of color under different illumination conditions, such as in bright daylight

versus overcast days). Neurons in ventral stream areas also display selective responses to naturalistic texture stimuli (Okazawa *et al.*, 2015); neuronal responses in earlier stages such as V1 and V2 largely represent information about low-level image statistics, while in later stages these responses reflect the material properties of the object (Goda *et al.*, 2014).

*1.5 Feature selectivity in area V4*

Area V4 was first described as a brain region where neurons responded strongly and selectively to colors of stimuli (Zeki, 1973, 1980)**.** Later characterizations revealed that responses of V4 neurons also reflect aspects of stimulus form (Desimone and Schein, 1987; Gallant *et al.*, 1996; Pasupathy and Connor, 1999). Pasupathy and Connor (2001) were able to capture shape-selective responses of V4 neurons as a function of the stimulus boundary curvature at a particular angular position with respect to the object center; that is, V4 neurons reflect information about partial object boundaries in an object-centered reference frame. Additionally, neurons in this area maintain their selectivity within the receptive field, *i.e.* they display translation-invariance, and some display size-invariant responses (Rust and DiCarlo, 2010; El-Shamayleh and Pasupathy, 2016). Since humans readily recognize objects at different locations and distances in the world, these properties of responses in area V4 may be important for enabling identification in real-world conditions.

Importantly, while previous studies have captured many of the response properties of V4 neurons such as their form selectivity, it is still unclear how this information is

represented jointly with other object attributes such as color and texture, and how the diverse response properties of the V4 population are decoded by downstream areas.

*1.6 Area V4: mid-level visual object representation*

Although feature selectivity has been studied extensively in the ventral visual stream, the field remains limited in its understanding of how information about different features ultimately leads to holistic perception and enables visual behaviors such as object recognition.

Due both to its anatomical location and the response properties of its neurons, visual area V4 is a particularly attractive area to deeply investigate the details of how the primate brain recognizes objects. V4 is an intermediate stage in the ventral stream of visual processing; anatomically, it is situated on the anterior bank of the lunate sulcus, and extends ventrally into the inferior occipital sulcus (Zeki, 1973). This area receives feedforward connections primarily from areas V1 and V2, and projects to the posterior portion of IT cortex. Understanding how visual processing occurs at this stage is thus beneficial to uncovering how information is combined from earlier areas (V1/V2), and how it informs downstream computations in IT and beyond. Importantly, lesions in area V4 are known to produce a deficit in recognition tasks, suggesting that responses of its neurons play a critical role in object representation (Schiller and Lee, 1991; Schiller, 1995).

*1.7 Influence of cognitive factors on responses in area V4*

In addition to its place in the middle of the visual object processing pathway, area V4 is well-connected with downstream areas involved in memory and learning.

Responses in V4 are modulated when the animal is focusing on a particular location or part of an object compared to when he is focusing elsewhere, even when the stimulus within the receptive field remains the same. This response property has been reported by a multitude of studies investigating the influence of attention to spatial location or object features on responses in V4 in the past 30 years (Moran and Desimone, 1985; Haenny and Schiller, 1988; Motter, 1993; Luck *et al.*, 1997; Reynolds *et al.*, 1999, 2000; McAdams and Maunsell, 1999, 2000; Maunsell and Treue, 2006) and makes area V4 an attractive candidate area at the intersection of perception and action. Our understanding of the influence of behavior on neuronal encoding is limited, especially with regard to how task-relevant information may be extracted from sensory responses. Thus, the effect of behavioral demands on responses in brain regions such as V4 is an important area for ongoing investigation.

### *1.8 Dissertation work scope*

In this chapter, I have provided an overview of selected literature relevant to our current understanding of object representation in the primate brain. Broadly, the goal of this dissertation is to bridge several disparate lines of investigation of how the brain represents objects, specifically focusing on the responses of neurons in area V4:

Aim 1. To investigate how V4 neurons represent information about object form and surface properties.

Many studies of feature selectivity investigate neuronal responses to one aspect of visual objects, *e.g.* form (Desimone and Schein, 1987; Kobatake and Tanaka, 1994;

Gallant *et al.,* 1996; Pasupathy and Connor, 1999, 2001) or color (Schein and Desimone, 1990; Conway *et al.*, 2007); studies of joint representation have been few and typically limited by experimental constraints (Bushnell and Pasupathy, 2012; Kobatake and Tanaka, 1994). V4 neurons display shape-selective responses, but is unknown to what extent these responses are influenced by surface characteristics of stimuli, such as interior fill.

**In Chapter 2, I will address this aim by exploring whether interior fill contributes to shape-selective responses in area V4.** A current best computational model of V4 selectivity predicted that, since signals in V4 reflect object boundaries, they should also be fill-invariant. I will demonstrate that only a small proportion of V4 neurons match this prediction, and in fact responses of most neurons in V4 are modulated by object fill in addition to object form. I will discuss the modifications necessary for the model to capture this joint representation of form and surface, reflecting the information from upstream visual areas that must be available for computations occurring in area V4 and beyond.

<u>Aim 2. To investigate whether object feature representations change with behavioral context.</u>

In addition to selectively representing several aspects of objects, V4 responses are influenced by cognitive factors. Is a neuron's feature selectivity maintained when the animal performs different tasks, or can feature tuning change as a result of behavioral engagement? Furthermore, do tuning changes reflect consistency with behavioral

demands, *i.e.* will a neuron's responses reflect a change in shape selectivity while an animal is using information about shape?

**In Chapter 3, I will address this aim by exploring whether feature selectivity changes when an animal performs tasks with different levels of behavioral engagement: a passive fixation task, and an active discrimination task.** I will present a comparison of single neuron responses to identical stimuli presented while the animals perform these two tasks. I will demonstrate that responses of a majority of neurons in area V4 are modulated in a stimulus feature-dependent manner; the rest show modulation that is dependent on firing rate, which has been reported previously (McAdams and Maunsell, 2000). I will then discuss how the nature of the feature-dependent modulation relates to the behavioral demands of the task.

Finally, in Chapter 4, I will provide a synopsis of the findings of my dissertation work and some perspectives on future investigations that could continue to improve our understanding of how responses in area V4 underlie object perception and recognition.

# Chapter 2. Aim 1: Representation of boundary shape and interior fill in area V4

## 2.1 INTRODUCTION

### *2.1.1 Representation of form in area V4*

Area V4 is an important stage in the cortical processing of visual form (Heywood and Cowey, 1987; Schiller and Lee 1991; Schiller, 1995; Merigan and Pham, 1998; De Weerd *et al.*, 1996; Merigan, 1996; Gallant *et al*., 2000). V4 neurons are responsive to simple shape features that are more elaborate than the edges and corners which elicit robust responses in neurons in areas V1 and V2; in particular, responses of V4 neurons appear to be sensitive to the conformation of object boundaries (Kobatake and Tanaka, 1994; Pasupathy and Connor, 1999; 2001). Since object boundaries are a reliable feature for identification, such a representational basis can directly support object recognition: if V4 neurons reliably signal information about object boundaries, it should be possible to decode object identity from their responses.

However, responses of V4 neurons are also sensitive to properties of objects that are related to their surface characteristics, such as color, luminance, and texture. In addition to shape, these aspects of the object reflect its identity and are important for recognition in natural scenes, since they can vary depending on the scene arrangement (foreground/background). Tuning for conformation of boundaries in V4 has been demonstrated largely with visual stimuli defined both by an outline and an interior fill. As a result, we do not know the relative contribution of object boundaries and interiors to the responses of shape-selective V4 neurons. Neurons in both up- and down-stream areas

display responses that are sensitive to the contrast polarity of stimuli, a property related to object interior (V2: Zhou *et al.*, 2000; IT cortex: Ito *et al*. 1994), but such stimuli have not been extensively tested in area V4.

### *2.1.2 Intersection with models of ventral stream processing*

Learning how object boundaries and interiors contribute to responses of V4 neurons is critical for understanding how V1 and V2 signals are combined to give rise to form selectivity in V4. In particular, a prominent model for V4 form selectivity proposed by Cadieu and colleagues (2007) implicitly assumes that form selectivity in V4 is based solely on boundary form. In this instantiation of the hierarchical-max (HMax) model, signals in the earliest stages are pooled across location and phase to capture key aspects of V4 form selectivity and invariance. This phase-pooling operation implies that the later stages of the model ignore interior fill and thus respond similarly to filled shapes and their outlines. This prediction suggests that if V4 neurons display similar responses to filled and outline stimuli, it is likely that they also discard phase information present in signals arriving from V1 and V2.

### *2.1.3 Chapter overview*

The following chapter will address whether boundary-based form selectivity in V4 is also influenced by surface characteristics. To answer this question, we studied the responses of 43 V4 neurons in two fixating macaque monkeys to a set of filled and outline

stimuli (Pasupathy and Connor, 2001). These stimuli were identical in their boundary conformation, but differed in the interior fill; the HMax model predicted similar responses to the stimuli that shared the same boundary, regardless of the interior. A few neurons responded similarly to filled and outline stimuli, consistent with the HMax prediction. However, responses of the majority of neurons were influenced both by boundary features and interior fill. To better explain how V4 responses arise, we implemented the HMax model and tested two strategies for retaining surface information within the model—eliminating phase-pooling and including unoriented channels that conveyed sensitivity to surface fill. Both modifications successfully produced model V4 neurons that responded to filled and outline stimuli.

## 2.2 METHODS

### *2.2.1 Surgical methods*

We implanted 2 adult male macaque monkeys (*Macaca mulatta*) with custom headposts and chambers positioned over dorsal area V4 (left hemisphere). The placement of the chamber over the prelunate gyrus and subsequent craniotomy were guided using structural MRI (for details, see Bushnell *et al.,* 2011a). All animal procedures conformed to the National Institutes of Health guidelines and were approved by the Institutional Animal Care and Use Committee at the University of Washington.

### *2.2.2 Visual stimulus presentation*

Stimuli were presented against a uniform gray background (luminance 5.4 cd/$m^2$) using a spectrally calibrated (PR650, PhotoResearch) CRT monitor positioned 45.4 or 56 cm away (subjects 1 and 2, respectively). The animal fixated a small white dot in the center of the screen within a window of radius 0.75°-1°. Eye position was monitored using an infrared eye tracking system (EyeLink 1000; SR Research) and coordinated with stimulus presentation using custom software based on Pype (developed by James Mazer). The animal was rewarded with drops of juice for successful fixation for the entire duration of a trial (typically 2000-2500 ms).

*2.2.3 Stimulus design*

We used a set of 51 shapes (**Fig. 2.1A**) presented at 8 rotations at 45° increments to assess neuronal shape selectivity. The design of these stimuli is described in detail elsewhere (Pasupathy and Connor, 2001). Stimuli were positioned in the center of a neuron's RF and scaled such that all parts of the stimuli were within 80% of the RF diameter, which was estimated as a function of eccentricity based on data in Gattass *et al.* (1988). Stimuli were offset from the center and rotated, so different rotations of each stimulus sampled different spatial locations within the receptive field (RF). Stimuli with rotational symmetry (*e.g.* circles) were treated as identical across congruent rotations (*i.e.* responses were averaged), producing a total of 362 unique stimuli in the set.

Outline stimuli were constructed using the same boundary form, but with an interior fill equivalent to the background (**Fig. 2.1B**). Outline width varied from 2 to 4 pixels, with thicker outlines for neurons with more eccentric RFs. Stimuli were presented for 300 ms; 4-5 stimuli were typically shown per trial, with a 200 ms blank interval preceding each stimulus. Filled and outline stimuli were randomly interleaved and presented multiple times. We include neurons for which we were able to repeat each stimulus at least 3 times (median: 5 repetitions). To evaluate whether differences in neuronal responses to filled and outline stimuli were attributable to overall stimulus intensity, for 20 neurons we studied the responses to a subset of 6-8 outline shapes with outline widths ranging from 1-6 pixels (see Results).

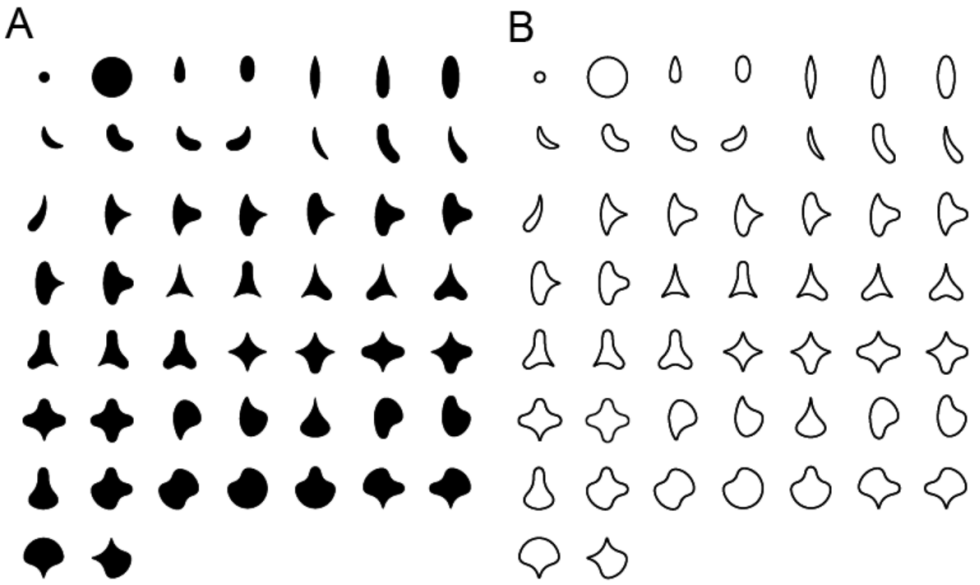


**Figure 2.1**. *Filled and outline stimulus set*.

Filled (A) and outline (B) stimuli used to study responses of V4 neurons. Stimuli were presented in up to 8 rotations (fewer in case of rotational symmetry) in 45° increments. All stimuli were scaled to fit within the RF and were presented in a preferred color and luminance for the neuron under study (see Methods for details).

### *2.2.4 Data collection*

Each day, we lowered a single tungsten microelectrode (FHC) into area V4 using a stepper motor drive (Gray Matter Research). We amplified and filtered the signals from the electrode and sorted waveforms (Plexon Systems) to identify single units. In this study, we include responses from 43 well-isolated neurons for which RF eccentricities ranged from 0.4° to 6.1°. After assessing the spatial extent of the receptive field manually, we characterized each neuron's color/luminance preference either manually or via an automated protocol systematically sampling responses to 25 colors in CIE space at 3-4 luminance levels (see Bushnell *et al*., 2011b). All stimuli were subsequently presented in a color and luminance that elicited a robust response from the neuron.

*2.2.5 Analysis*

Spiking responses were averaged across multiple presentations of the same stimulus and over the 300 ms duration of each stimulus presentation. Stimulus onset and offset times were detected with a photodiode.

We computed two metrics to compare the responses to filled and outline stimuli. First, to assess similarity of tuning, we computed Pearson's correlation coefficient, r, between responses to filled and outline stimuli for each neuron. To compare the range of responses evoked by filled stimuli to that evoked by outline stimuli, we computed a relative response range metric (*ρ*) for each neuron. For this, we first computed the average absolute deviation (AAD) from the mean for responses to filled or outline stimuli as,

$$\mathrm{AAD} = \frac{1}{n}\sum |\bar{x_i} - \bar{x}|$$

*Equation 2.1*

where $\bar{x}$ is the mean response of the neuron to a stimulus type (*i.e.* filled or outline), $\bar{x_i}$ is the mean response to a unique stimulus across all presentations, and $n$ = 362 stimuli. We then computed the relative range metric as follows:

$$\rho = \frac{AAD_{outline}}{AAD_{filled}}$$

*Equation 2.2*

A *ρ* value near 1 implies similar responsiveness to filled and outline stimuli; *ρ* > 1 implies a larger range of responses to outline than to filled stimuli, and the converse for *ρ* < 1.

We obtained identical results using standard deviation to quantify response spread instead of AADM.

To quantify the overall luminance associated with each stimulus, we computed a stimulus intensity metric by dividing the number of pixels used to generate each shape by the number needed to generate the largest filled shape (the large filled circle, **Fig. 2.1A**). Values ranged from 0.026 to 0.19 for outline shapes, and 0.068 to 1.0 for filled shapes.

To characterize the shape selectivity of each neuron in terms of tuning for boundary curvature, we fit responses to the angular-position curvature (APC) model of Pasupathy and Connor (2001). Briefly, we represented each stimulus in our set in terms of the curvature and angular position of 8 contour segments along the boundary of each shape. Curvature values ranged from -0.3 (moderate concavity) to 1.0 (sharp convexity); angular position values ranged from 0° (right of center) to 360° in a counter-clockwise direction. Using non-linear regression, we identified the best-fitting 2D Gaussian in a space with dimensions of angular position and curvature that best predicted observed responses. The 2D Gaussian function was defined by two means ($\mu_{curv}$, $\mu_{\theta}$), two standard deviations ($\sigma_{curv}$, $\sigma_{\theta}$), for the curvature and angular position dimensions, respectively, and an amplitude, that best predicted the observed responses.

*2.2.6 Model architecture and function*

2.2.6.1 Overview. Our instantiation of the feedforward hierarchical-max (HMax) model follows closely that of Cadieu *et al.* (2007) and has four main stages. The first two stages are named S1 and C1, in an analogy to simple and complex cells in the primary visual

cortex (V1). The S1 units have oriented Gabor function linear filters (**Fig. 2.2A**, icons at bottom) that operate on the input stimulus. The C1 units then compute a maximum (Max) function over many shifted S1 units (**Fig. 2.2A**, triangle) that have the same orientation, thereby creating phase- and position-invariant tuning for orientation analogous to that of V1 complex cells. The third stage, S2, contains units that compute a weighted average (**Fig. 2.2A**) of the outputs of a particular set of C1 units. An S2 unit can be thought of as defining a flexible shape template based on the positions and orientations of its chosen C1 inputs. The template is flexible because of the position-invariance built into the C1 units. The final stage, C2 (**Fig. 2.2B**), consists of a Max function over a grid of spatially offset, but otherwise identical, S2 units. This is designed to achieve greater translation invariance. This model was designed to respond selectively to particular shapes regardless of their position within the C2 receptive field, thereby matching the boundary curvature selectivity and translation invariance of V4 neurons (Cadieu *et al*., 2007). Each stage in the model is described in detail below, and possible differences from the implementation of Cadieu *et al*. are noted.

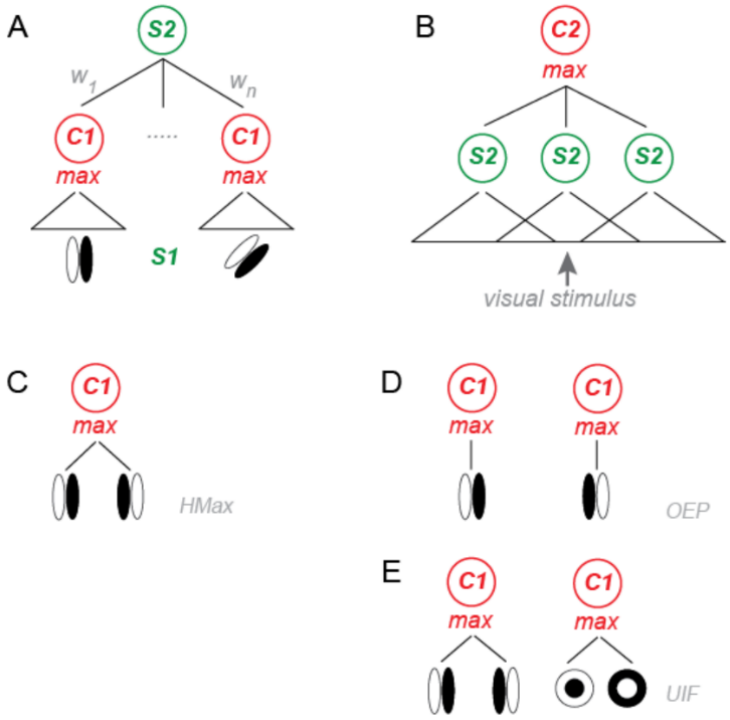


**Figure 2.2**. *HMax model structure and proposed modifications*.

(A-B) The HMax model described by Cadieu *et al*. (2007) creates V4-like units selective for specific boundary orientation through alternating operations imparting selectivity (S layers) and invariance (C layers).

(A) S1 units (icons) are Gabor filters varying in position, orientation (4 orientations), phase (2 phases), and size (6 sizes). C1 units are arranged in 3 spatial grids (2x2, 3x3, 4x4), each receiving inputs from oriented S1 units and taking a maximum (triangle) response over S1 unit phase, position, and size (2 sizes per grid, inversely proportional to grid tiling). Each C1 grid location has 4 C1 units, one per orientation. Each S2 unit receives input from 2-25 C1 units. The selection and strength of these weights ($w_1$ to $w_n$) are determined by the fitting procedure to produce a shape-selective template matching the response profile of a given V4 neuron. See Methods for implementation details.

(B). The C2 unit produces a V4-like response by taking the maximum response over a 3x3 grid of S2 units.
(C-E). Comparison of original HMax model and the proposed modifications.

(C). In the HMax model, each C1 unit takes a maximum response over S1 unit phases in addition to spatial position and size.

(D). To preserve information about interior fill, in the oriented edge polarity (OEP) model, each C1 unit takes a maximum response over position and size, but not phase of S1 units. This results in a doubling of the number of C1 units available for fitting.

(E). As an alternative strategy, in the Unoriented interior fill (UIF) model, interior fill information is made available to the S2 stage by adding unoriented difference-of-Gaussian units in the S1 stage. Each C1 unit takes a maximum response over position, size, and phase for oriented and unoriented units. This add 29 additional units to the C1 stage.

Visual stimuli presented to the model were represented as grayscale images on a 256 x 256 pixel grid; the entire field was a 12.8° x 12.8° square patch. Constants were set to be consistent with the original paper (Cadieu *et al*., 2007).

2.2.6.2 S1 stage. This stage consists of Gabor (sinusoidal function * Gaussian) linear filters at four orientations (0°, 45°, 90° and 135°), two spatial phases (either 0° and 180°, or 90° and 270°), and six spatial scales. Each filter is represented on a square 'tile' of NxN pixels where the tile size, N, defines the scale and is an element of {32, 36, 40, 45, 54, 60}. The spatial frequency (SF) of the Gabor filter is inversely proportional to N,

$$SF = \frac{3}{N}$$

*Equation 2.3*

in cycles/pix and the SDs of the spatial Gaussian (orthogonal and parallel to the sinusoid) are, in pixels,

$$SD_{orth} = \frac{N}{9}$$

*Equation 2.4*

$$SD_{par} = \frac{N}{6}$$

*Equation 2.5*

To establish units comparable to electrophysiology, we define 1 pixel to be 0.05° of visual angle. Thus, the smallest filters (tile size 32) have SF = 1.875 cyc/° and $SD_{orth}$ = 0.18°, and the largest filters have SF = 1 cyc/°and $SD_{orth}$ = 0.33°.

The set of S1 responses for each image is computed as the dot product between the filter and the image at each location in the image. The response is then normalized as,

$$r_{norm} = \frac{r_{raw}}{\sqrt{\sum_{i,j} stim^2} + \varepsilon_{s1}}$$

*Equation 2.6*

where the $\Sigma$ is over a local region (covered by the filter at location *i,j)* that scales with the tile size. In particular, it consists of all pixels in the stimulus that are within the radius $SD_{par}$ * 2.5. The region used in the original Cadieu *et al.* study is unspecified. The constant, $\varepsilon_{s1}$, is set to 0.001 to avoid division by zero.

2.2.6.3. C1 stage. The C1 units tile the input visual grid overlapping S1 units, and are arranged in three grids: a 4x4 grid with smaller RFs, a 3x3 grid with medium RFs and a 2x2 grid with larger RFs. The spacing between C1 units at neighboring grid positions is 24, 30 and 40 pix, respectively (see Table 1 in Cadieu *et al*., 2007). Each C1 unit takes the Max over S1 units at two spatial scales and at a specific orientation preference. For example, the C1 units in the 4x4 grid take the Max over all S1 units at the first and second tile sizes, *i.e.*, N=32 and 36 (see S1 stage, above) that have a given orientation preference and that have their RFs contained within the 48 pix square centered on the C1 unit grid position. The Max is also taken over filters having two spatial phases that are 180° apart (**Fig. 2.2C**). Thus, at each grid position, there are four C1 units corresponding to the four orientations of the S1 Gabor filters. The C1 units in the 3x3 and 2x2 grids operate similarly, except their RF limits are 60 and 80 pix, respectively (rather than 48 pix), and their S1 inputs come from tile sizes 40 and 45 (3x3 grid) and 54 and 60 (2x2

grid). Overall, the three grids, with 4 orientations at each point, provide 116 C1 units (varying in position, orientation and scale) for the next stage, S2.

2.2.6.4. S2 stage. Each S2 unit computes a weighted sum of a set of C1 inputs (**Fig. 2.2A**) as follows,

$$u = \frac{\sum w_i x_i}{\sqrt{\sum x_i^2 + \varepsilon_{s2}}}$$

*Equation 2.7*

where $x_i$ is the response of the *i*th C1 unit, $w_i$ are the weights and $\varepsilon_{s2}$ = 0.0001. The denominator normalizes the response in a way that makes *u* relatively insensitive to a scaling over all input signals. The normalized, weighted sum is then passed through a sigmoid function:

$$g(u) = \frac{s}{1 + e^{-a(u-b)}}$$

*Equation 2.8*

where *s* is the maximum value, *a* is the slope, and *b* is the horizontal offset. These three sigmoid parameters and the weights in the previous equation are the only parameters that are fit to the V4 data (see *Model fitting* below).

S2 units were arranged in a 3 x 3 grid, with a spatial offset of 30 pix between neighboring units. All nine S2 units have identical weights on their C1 inputs, and each S2 unit receives inputs from a set of C1 grids that are centered on its S2 grid-point.

2.2.6.5 C2 stage. There is one C2 unit, and it computes Max over the 3 x 3 grid of S2 units. This C2 unit thus has a larger RF than the S2 units, and its output is the value compared to V4 responses during training and testing.

### *2.2.7 Model fitting and testing*

The HMax model was fit to the V4 neuron responses in a manner similar to that described by Cadieu *et al*. The neuronal data consisted of the responses to the 362 shape stimuli. These response were divided randomly into a training set (5/6 of the data) and a testing set (1/6) of the data. During the training phase, only the three parameters of the S2 sigmoid (s, a, and b, *Eqn. 2.8*) and the weights from the C1 to the S2 layers were varied. Unlike Cadieu *et al.*, we allowed the model to assign positive as well as negative weights for C1-S2 connections to expand the parameter search space to excitatory/inhibitory influences. The training phase consisted of two parts: first, the best fit was found for a model having only two non-zero weights from the C1 to the S2 layer, a procedure replicating the approach used by Cadieu *et al.* In particular, all unique pairs of C1 weights were tried from the middle C1 grid (the 3x3 grid with 4 orientations at each point). For each pair of weights, gradient descent was performed on the five parameters (s, a, b, w1, w2) from many starts (typically 625). Once the best model was found that had two C1 units from the middle grid, the second phase of fitting began. Here, we iteratively tested and included the next best C1 unit by trying all remaining C1 units and fitting a weight, $w_i$, along with the three sigmoid parameters, using gradient descent from

an array of starting points for the weight of the additional C1 unit. We continued this procedure until 25 C1 units had been chosen.

For each of the 43 neurons described here, we repeated this fitting procedure 10 times, using 5/6 of the responses (randomly selected) as the training set and the remaining 1/6 as the testing set. We computed a mean squared error (MSE) across the 10 test sets as a function of the number of additional C1 units selected by the fitting procedure. We then determined the smallest number of C1 units necessary for the model to achieve an MSE value within 1% of the minimum average cross-validated MSE; we restricted this number between 2-25, as in Cadieu *et al*. We obtained final results for each neuron by fitting the model given all neuronal data (combining the training and testing set) and constrained to select a certain number of C1 units as determined above.

We compared model responses to neural data using Pearson's correlation coefficient (r) between predicted and observed responses, the mean squared error (MSE) from the fitting procedure, and the relative response range metric ($\rho$) described above.

*2.2.8 Model modifications*

To better describe the responses of neurons which were dictated by interior fill in addition to boundary form, we considered two modified versions of the HMax model which were targeted towards capturing the differences between filled and outline stimuli at the earliest stages of the network. We considered both a model where V4-like model units received information about interior fill from the same early-level units that provided

information about the oriented boundary, and another where information about interior fill was separated as a separate population of unoriented early-level units:

Oriented edge polarity (OEP). We provided the model with access to phase information present in S1 level Gabor filter units. The original HMax model throws away information about spatial phase, and thus the polarity of edges, by taking the Max over 2 complementary phases (**Fig. 2.2C**). We thus created a variant of the model that maintains this phase information, passing it from the S1 to C1 level (**Fig. 2.2D**), and we refer to this as the oriented edge polarity (OEP) model. This change doubled the number of C1 units (116 for each phase), for a total of 232 units available during the selection of the best 2-25 C1 inputs to each S2 unit, but the rest of the model structure and fitting procedure were preserved.

Unoriented interior fill (UIF). To evaluate whether unoriented units, separate from those representing the oriented boundary, were important to signal object fill, we added difference-of-Gaussian units to the pool of oriented Gabors in the S1 layer (**Fig. 2.2E**), within the original framework in which C1 units max over phase. We refer to this model variant as the unoriented interior fill (UIF) model. The Gabor RFs for the unoriented units were generated as follows:

$$DoG = e^{\frac{-((x-x_0)^2+(y-y_0)^2)}{2{\sigma_{center}}^2}} - e^{\frac{-((x-x_0)^2+(y-y_0)^2)}{2{\sigma_{surround}}^2}}$$

*Equation 2.9*

where *x* and *y* are coordinates along the horizontal and vertical axes of space, $x_0$ and $y_0$ define the spatial center of the DoG, $\sigma_{center}$ and $\sigma_{surround}$ are the standard deviations of the central and surround Gaussian components, respectively (N/5 and N/3, where N is

the tile size defined in the section “S1 stage” above). Gaussian components were normalized before subtracting.

This model modification added 29 unoriented C1 units to the original 116 oriented C1 units, for a total of 145 units available during the selection C1 to S2 inputs. Other than these variations, the rest of the model architecture and fitting procedure remained identical.

## 2.3 RESULTS

Below, I first describe our experimental observations related to the neuronal encoding of filled *vs*. outline shapes in V4, and then demonstrate how a biologically inspired hierarchical-max (HMax) model, previously used to explain V4 shape tuning, can be altered to account for our new data.

### *2.3.1 Responses of V4 neurons to filled and outline shapes*

To determine whether form selectivity in area V4 is invariant across filled and outline shapes, we recorded responses from well-isolated V4 neurons (n = 43) in two awake, fixating macaques for the two stimulus sets shown in Figure 2.1**.** The filled shapes (**Fig. 2.1A**) consisted of unique combinations of convex and concave features and have been used in a series of studies to characterize V4 tuning (Pasupathy and Connor, 2001; Bushnell *et al.*, 2011a; Bushnell *et al*., 2012; Oleskiw *et al*., 2014; Kosai *et al*., 2014). In comparison, the outline shapes had the same boundary curvature and were rendered in the same color and luminance as the filled shapes.

#### *2.3.1.1 Example neurons*

Neurons in V4 showed a variety of responses to filled and outline shapes, which we demonstrate with four example neurons. Neuron 1 (**Fig. 2.3**) exemplifies the responses of those V4 neurons that exhibited similar shape preferences for filled and

outline stimuli. This neuron responded strongly to many of the filled shapes that included a concavity to the right relative to the center of the shape (**Fig. 2.3A**, red boxes). In comparison, stimuli with a convexity to the right of the shape evoked virtually no spikes (**Fig. 2.3A**, green boxes). This neuron's responses to outlines (**Fig. 2.3B**) were weaker than those to filled shapes, but the shape tuning was similar: responses were stronger for outlines with a concavity to the right and weaker for outlines with a convexity to the right of object center.

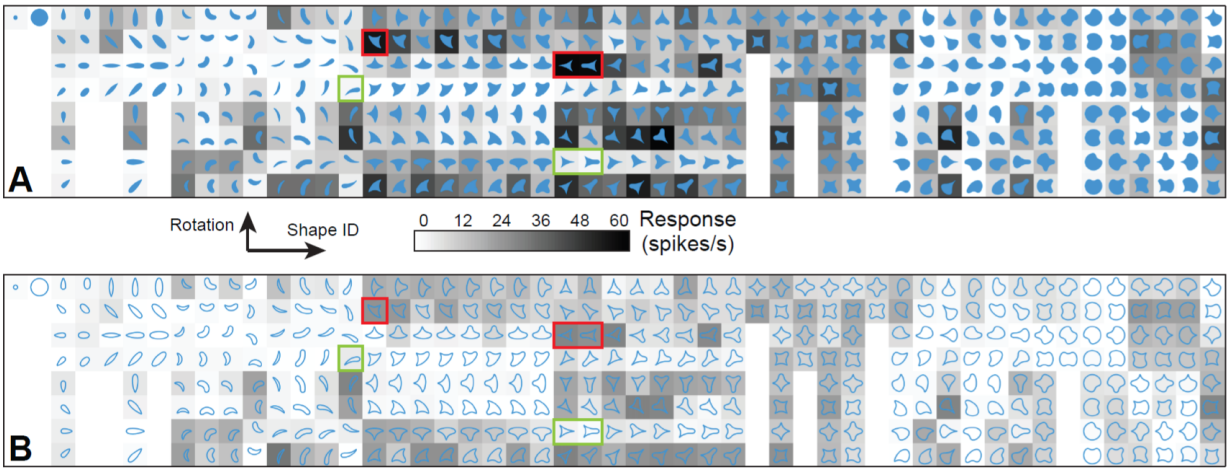


**Figure 2.3**. *Responses of an example V4 neuron to filled and outline stimuli*.

Responses of Neuron #1 (o140915) to filled (A) and outline (B) versions of 51 shapes (abscissa) at up to 8 rotations (ordinate) are shown. The background gray level of each icon represents the mean response to the stimulus drawn within. Darker backgrounds correspond to higher firing rates (see scalebar).

Responses were weaker for outline stimuli, but shape preferences were similar. Stimuli that elicited stronger spiking responses from this neuron (red boxes) included a concave feature to the right side of the shape, while shapes that evoked a weaker response (green boxes) did not.

Responses to both filled and outline stimuli were well-fit by an angular position and curvature (APC) model with a shallow concavity to the right. Fitted means based on filled stimuli: $\mu_\theta$ = 350°; $\mu_{curv}$ = -0.12; outline stimuli: $\mu_\theta$ = 345°; $\mu_{curv}$ = -0.14. In both cases, correlation between predicted and observed responses was 0.87.

Responses to filled and outline stimuli for each shape were highly correlated (**Figure 2.4A;** $r = 0.91$, $p < 0.001$) indicate shape tuning was similar for both sets of stimuli. We used the relative response range, $\rho$ (see Methods), to quantify which stimulus category (filled or outline) elicited relatively more selective responses from the neuron, since a narrower range of responses to one category of stimuli reflects lower selectivity (modulation ranging from best to worst stimulus) for that category. The $\rho$ value for responses of Neuron 1 was 0.59, indicating a narrower range of responses to the outline stimuli and reflecting this neuron's stronger responses to filled stimuli. In summary, both sets of stimuli modulated the responses of this neuron, and shape preferences were similar regardless of object fill.

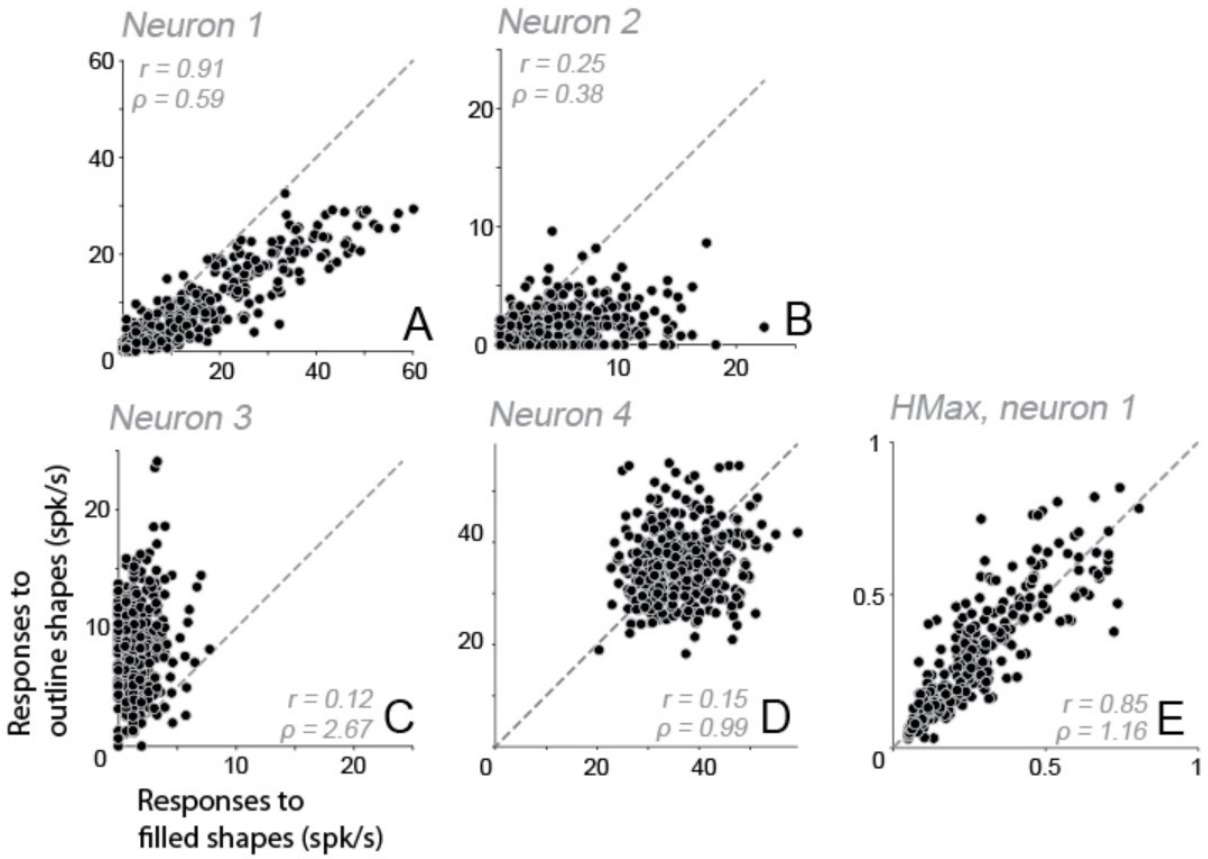


**Figure 2.4**. *Responses of example V4 neurons and* HMax *model predictions*.

(A-D). Responses of neurons 1-4 to filled and outline stimuli. Each point represents the mean response (spk/s) to the presentation of an outline (ordinate) and a filled (abscissa) stimulus with the same boundary shape.

(A) For Neuron #1 (same as in **Fig. 2.3**), responses to filled and outline stimuli were strongly correlated ($r = 0.91$, $p < 0.001$). Filled stimuli evoked a larger range of responses than outline stimuli and the relative range of responses ($\rho$) was 0.59.
(B). For Neuron #2, filled shapes evoked a large range of responses than outline shapes. Correlation between responses to filled and outline stimuli was weak ($r = 0.25$, $p < 0.001$) and $\rho$ was 0.38.
(C). For Neuron #3, outline stimuli evoked a large range of responses and $\rho$ was 2.67. Here again, correlation between responses to filled and outline stimuli was weak ($r = 0.12$, $p = 0.026$).
(D). Filled and outline stimuli evoked a similar range of responses from Neuron #4 ($\rho = 0.99$) but correlation between filled and outline responses was weak ($r = 0.15$, $p = 0.003$).
(E). Responses to filled and outline stimuli predicted by an instantiation of HMax model which was trained on responses of Neuron #1 to filled stimuli (abscissa in A). Most points lie along the line of equality; *i.e.* the model predicts similar responses to outline and filled shapes that share identical boundary configuration. Correlation between predicted responses to filled and outline stimuli was high ($r = 0.87$; $p < 0.001$) and $\rho$ was 1.16. The HMax model template included 25 C1 units and the predictions shown here are based on odd-symmetric S1 filters, which achieved better performance than even-symmetric filters in this case.

A strong correlation in responses between filled and outline stimuli, as observed in Neuron 1, would be expected from the HMax model (see Methods) for V4 shape selectivity proposed by Cadieu and colleagues (2007). In this model, shape selectivity and translation invariance are built by pooling weighted signals from V1-like oriented filters followed by Max pooling (**Fig. 2.3A, B**). Because the HMax model relies on bandpass orientation-tuned filters to build shape selectivity, we expected that it would predict a strong correlation between responses to filled and outline stimuli. To test this, we fit the HMax model to the filled responses of Neuron 1 (as done by Cadieu *et al*.; see Methods) and then used the model to predict outline responses. The best fit model (correlation between predicted and observed responses to filled stimuli = 0.76; $p<0.001$) did indeed produce a strong correlation ($r = 0.87$; $p<<0.0001$) between predicted responses to filled and outline shapes (**Fig. 2.4E**).

However, the model based on the responses of Neuron 1 produced a similar range of predicted responses to filled and outline stimuli ($\rho = 1.16$). A $\rho$ value close to 1 indicates that the model unit was driven similarly by filled and outline stimuli, unlike what we observed in the relatively stronger responses to filled stimuli for Neuron 1.

Many neurons in our dataset, however, were unlike Neuron 1: their responses to filled and outline stimuli were not well correlated. This difference is illustrated by example Neurons 2-4 in **Figure 2.4B-D**. Neuron 2 responded strongly and selectively to filled shapes (**Fig. 2.5A**), but its responses to outline shapes were much weaker and did not reflect the selectivity observed for filled shapes (**Fig. 2.5B**); this difference in responses was significant (paired t-test, $p < 0.001$). The scatterplot of outline *vs*. filled responses for Neuron 2 (**Fig. 2.4B**) was distinctly different from that for Neuron 1 (**Fig. 2.4A**). The responses to filled and outline shapes were poorly correlated ($r = 0.25$, $p<0.01$), and the relative response range, $\rho$, was 0.38, indicating that responses of Neuron 2 were better modulated by filled stimuli.

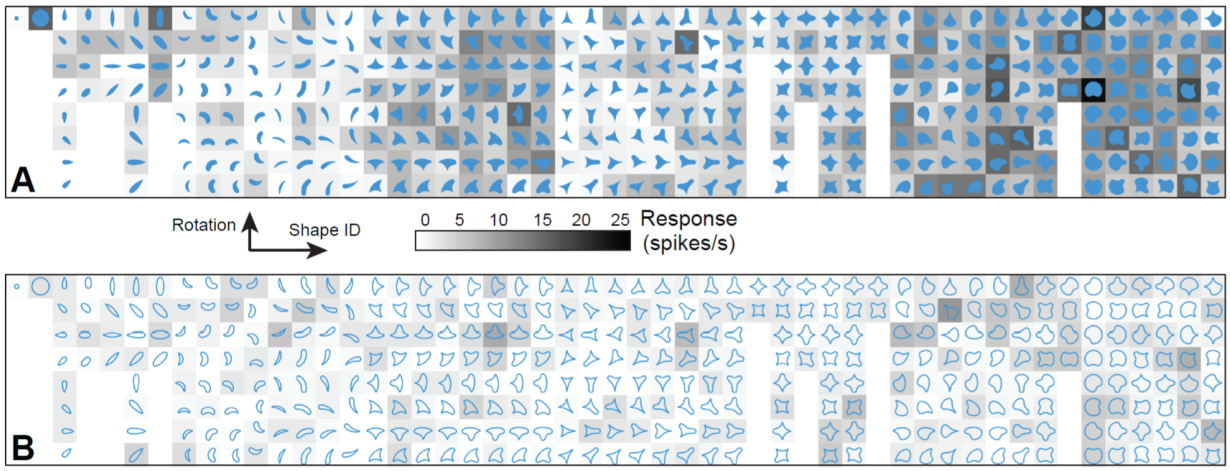


**Figure 2.5**. *Responses of example neuron #2 to filled and outline stimuli.*

Same format as in **Fig. 2.3**. Neuron #2 (p150428; also shown in Fig. 2.4B) responded strongly and selectively to filled stimuli, but not to outlines.

Neuron 3 responded strongly and selectively to outline stimuli (**Fig. 2.6B)**, but unlike the previous examples, it responded weakly to filled shapes (**Fig. 2.6A**); this difference in responses was significant (paired t-test, $p < 0.001$). There was little correlation between filled and outline responses ($r = 0.12$, $p=0.026$, **Fig. 2.4C**), and $\rho$ was 2.67, reflecting the greater range and stronger modulation in responses to outlines.

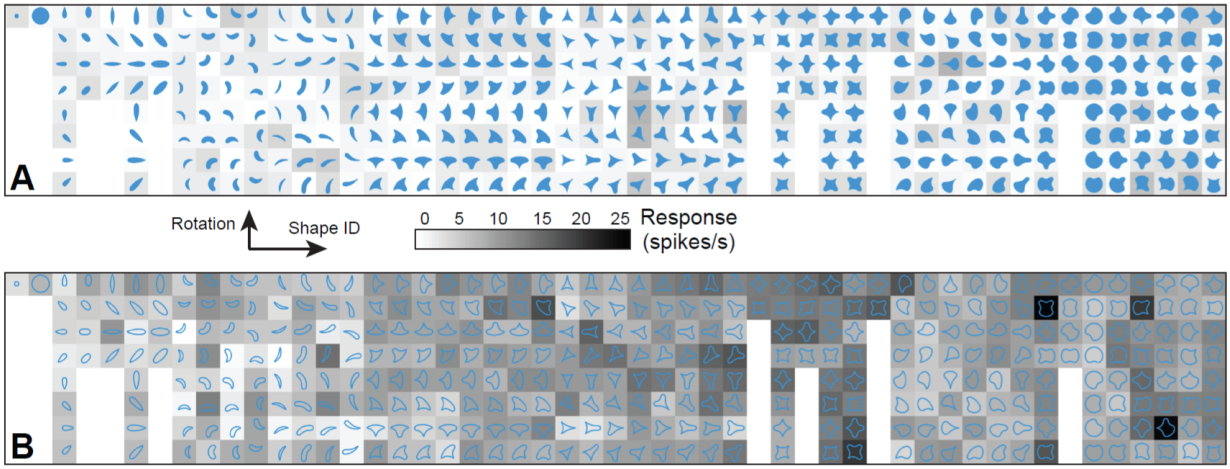


**Figure 2.6**. *Responses of example neuron #3 to filled and outline stimuli.*

Same format as in **Fig. 2.3**. Neuron #3 (p150415; also shown in Fig. 2.4C) responded strongly and selectively to outlines, but not to filled stimuli.

Neuron 4 (**Fig. 2.4D)** was driven well and over a similar range by both stimulus types ($\rho = 0.99$), but the patterns of selectivity for filled and outline shapes were dissimilar ($r = 0.15$, $p=0.003$). The diversity in these three examples, and their dissimilarity from the behavior of Neuron 1, can be visualized by comparing the four scatter plots for outline *vs*. filled responses depicted in **Figure 2.4A-D**. Only the behavior of Neuron 1 is roughly consistent with the predictions of the HMax model that responses for filled and outline stimuli should be similar in magnitude and strongly correlated.

*2.3.1.2 Population Data*

Across our dataset of 43 V4 neurons, we compared the distribution of the two response summary metrics: correlation in responses between filled and outline stimuli (**Fig. 2.7A**, abscissa) and the relative range of responses $\rho$ (ordinate). While a few of the neurons we recorded showed a strong correlation between responses to filled and outline stimuli, a majority of the neurons did not: only 4 neurons had $r > 0.5$ and the median was 0.16. V4 neurons showed a broad range of $\rho$ values (0.2 to 2.67, median 0.69), with many neurons exhibiting a compressed range of responses to one or the other category of stimuli. The majority of neurons exhibited a larger range of responses across filled stimuli (33/43 neurons had $\rho < 1$), and the mean $\rho$ value was significantly less than 1 ($\mu = 0.86$; t-test using log-transformed $\rho$, $p < 0.001$).

For each neuron in our dataset, we fit the HMax model to the responses to filled shapes and predicted outline responses as before (*e.g.,* **Fig. 2.4E**). Across the 43 HMax models that best fit our data (one model per neuron), mean correlation and $\rho$ were 0.85 (SD = 0.11) and 1.35 (SD = 0.48) (denoted by x, **Fig. 2.7A**). These values deviated substantially from those observed across our dataset. Thus, our data suggest that V4 neurons show more diversity in their responses to filled and outline stimuli than the HMax model predicts.

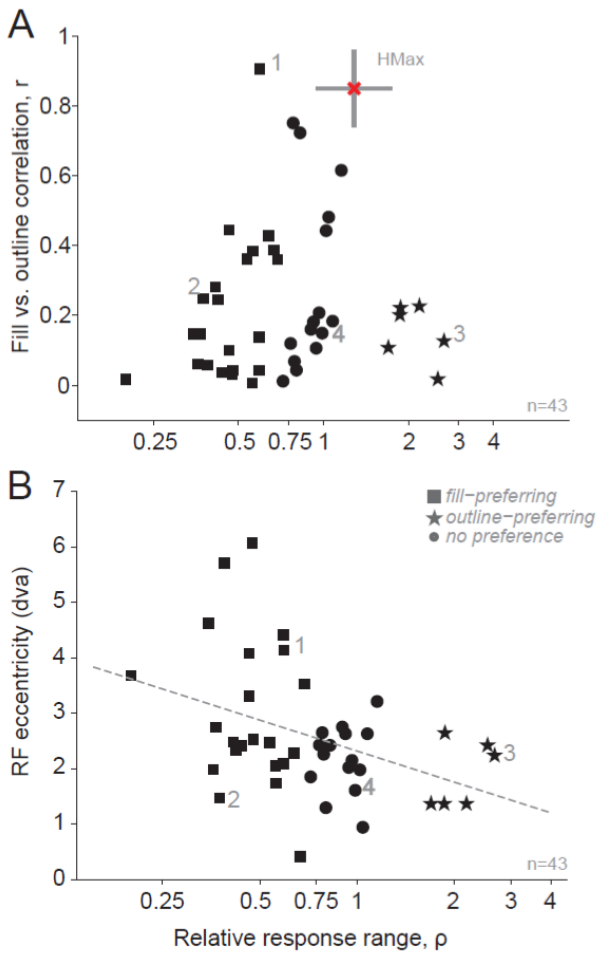


**Figure 2.7**. *Population responses to filled and outline stimuli*.

(A) Correlation between the responses to filled and outline stimuli (ordinate) are plotted against the relative range of responses to filled versus outline stimuli ($\rho$, abscissa; see Methods) for each neuron in our dataset. Neurons with $\rho > 1$ have a larger range in their responses to outline compared to filled stimuli, Neurons with $\rho > 1.43$ (stars) were classified as outline-preferring and neurons with $\rho < 0.7$ (squares) were classified as fill-preferring neurons. Mean across the recorded neurons for correlation and $\rho$ were 0.23 (SD = 0.22) and 0.86 (SD = 0.58) respectively. For comparison, mean (denoted by x) and standard deviation (grey bars) of fill/outline response correlation and $\rho$ across a set of 86 HMax models trained on data from the 43 V4 neurons (two models per neuron corresponding to even- and odd-symmetric S1 filters) are also shown. Example neurons 1-4 are identified.

(B) RF eccentricity (ordinate) is plotted against the relative range of responses to filled versus outline stimuli ($\rho$, abscissa) for each neuron. There was statistically significant negative correlation ($r = -0.4$, $p = 0.008$), suggesting that filled stimuli evoked a larger range of responses at larger eccentricities. Dashed line indicates the regression slope.

Prior to developing models that capture our experimental observations (below), we explore whether the diversity in our metrics across neurons was associated with variations in other characteristics that provide insight into how they could be modeled, including properties such as their receptive field (RF) eccentricity, preference for stimulus intensity, luminance contrast, and curvature tuning. We sub-divided neurons into 4 groups related to the four examples presented earlier (**Fig. 2.4**). Neurons with a fill-outline correlation greater than 0.4 ($p < 0.01$) were classified as fill-outline invariant neurons; this group included 8 of the 43 neurons (19%). In addition to displaying similar selectivity to

filled and outline stimuli, these neurons had $\rho$ values between 0.46 and 1.16 (median $\rho$ = 0.80)**.** Thus, responses of these 8 neurons were broadly consistent with the predictions of the HMax model. The remaining neurons with lower fill-outline response correlations (r < 0.4) were divided into three groups based on their relative range of responses. Specifically, if one category of stimuli (outline or filled) evoked responses with a range 40% greater than the other, it was considered to be the preferred stimulus category since it modulated neuronal responses more strongly**.** Thus, 6 neurons with $\rho$ >= 1.4 (stars, Fig. 2.7A) were classified as outline-preferring, and 19 neurons with $\rho$ < 0.7 (squares) were classified as fill-preferring. All outline-preferring neurons and 18/19 fill-preferring neurons exhibited significantly stronger responses to their preferred stimulus category (outline and filled, respectively; paired t-test, p < 0.001). The remaining 10 neurons (23%) with $\rho \approx 1$ (circles) and poor fill-outline correlation were typically driven well by most stimuli, and thus not shape selective; they exhibited low goodness-of-fit values from the APC model (median r = 0.3). While these category boundaries are arbitrary, it allowed us to ask if neurons with different patterns in their responses to filled and outline stimuli also differed in terms of other characteristics.

### *2.3.2 Factors contributing to fill selectivity*

*2.3.2.1* RF eccentricity. It is possible that the preference for filled versus outline stimuli is correlated with RF eccentricity. Filled and outline stimuli vary in terms of their spatial frequency (SF) content, with filled stimuli having more power at low SF. Given that SF tuning varies with eccentricity in V1 (Tootell, 1988) and that SF tuning properties in V4

are similar to those in V1 (Desimone and Schein, 1987), it is conceivable that neurons with more eccentric RFs may also respond better to filled stimuli. Across our population of neurons, RF eccentricity ranged from 0.4° to 6.1° (**Fig. 2.7B**, ordinate) and there was a statistically significant negative correlation between $\rho$ and RF eccentricity ($r = -0.4$, $p = 0.008$). For neurons with RFs within 4° of the fovea, $\rho$ ranged from 0.2 to 2.67, reflecting the presence of both fill- and outline-preferring neurons, as well as neurons that did not have a preference for one or the other category of stimuli. For neurons at more peripheral RF locations (eccentricity > 4°), $\rho$ was < 0.59 for all neurons, *i.e.* neurons responded more strongly to filled stimuli than to outlines. This bias toward fill-preferring cells at larger eccentricities is consistent with a preference for lower spatial frequencies away from the fovea.

*2.3.2.2.* Overall stimulus intensity. To determine whether preference for overall stimulus intensity could explain the differential responses to filled and outline stimuli, we studied the responses of 20 neurons with two outline widths, one twice the other (*e.g.* 2 and 4 pixels wide) for a subset of 5-6 shapes. Doubling outline width increased the amount of interior that was filled, and thus increased overall stimulus intensity relative to the background. We reasoned that, if stronger responses to filled stimuli were due to a preference for higher stimulus intensity, then regression slopes between filled and outline stimuli should be larger with increasing outline widths. For only 2/20 neurons, slopes for double outline width were significantly greater than for single outline width ($p < 0.05$). Of the remaining 18 neurons, none of the 8 neurons that preferred filled stimuli were associated with a significant increase in response due to doubling of the outline width. Across all neurons, the slope differences were not significantly different from 0 (paired t-

test, p = 0.33). Thus, doubling the width of outline stimuli did not appear to change responses, and preference for filled stimuli could not be simply attributed to overall stimulus intensity.

*2.3.2.3* Luminance contrast tuning. While a majority of V4 neurons are sensitive to stimulus luminance contrast, we previously found that a small fraction (~11%) are not (Bushnell *et al*., 2011b). The average luminance contrast of the stimulus relative to the background is different for outline and filled stimuli. Therefore, we wondered whether the insensitivity of some neurons to luminance contrasts relates to insensitivity to object fill. For six neurons classified as fill-outline invariant we characterized color and luminance preferences in detail (see Methods). Responses of 5/6 of these neurons were significantly modulated by luminance contrast (2-way ANOVA for luminance and color; significant main effect for luminance and/or interaction term, $p<0.05$) suggesting that insensitivity to luminance contrast is not a hallmark feature of neurons that respond similarly to filled and outline stimuli. Rather, such neurons can still be independently influenced by boundary shape and interior fill.

*2.3.2.4* Curvature tuning. To determine whether variations across V4 cells in their relative preferences for filled and outline stimuli related to their selectivity for boundary curvature, we fit the responses of our V4 neurons to the same angular-position and curvature (APC) model that we have used to characterize shape tuning in a series of previous studies (Pasupathy and Connor, 2001; El-Shamayleh and Pasupathy, 2016; Kosai *et al.*, 2014). For cells that responded similarly to filled and outline stimuli, curvature sensitivity assessed based on responses to filled and outline stimuli was similar: the median goodness-of-fit across 8 neurons was 0.5 and 0.48 respectively. For fill-preferring

neurons, the median goodness-of-fit coefficient for the responses to filled shapes was 0.47; thus, we found no difference in the curvature sensitivity of fill-preferring neurons and fill-outline invariant neurons. For outline-preferring neurons, the median goodness-of-fit coefficient for responses to outline stimuli was lower (r = 0.35), but this difference may be attributed to the low number of outline-preferring neurons we encountered—2 out of 6 such neurons had goodness-of-fit coefficient > 0.4. In summary, we were unable to find any clear relationship between boundary shape tuning and whether a neuron will have a particular preference for filled *vs*. outline shapes.

*2.3.3 Modelling improvements*

Results presented above suggest that the HMax model, previously shown to be a good fit to V4 boundary curvature tuning (Cadieu *et al*., 2007), predicts similar responses for filled and outline shapes and thus cannot capture the diversity in our V4 data. Our evaluation above was based on fitting the HMax model with filled stimuli alone. To more comprehensively assess the ability of the HMax model to capture a range of fill-outline correlations and $\rho$ values, we refit the model using neuronal responses to both filled and outline stimuli and assessed correlation and $\rho$ values from the predicted responses. Training on both filled and outline stimuli did not produce a model that matched typical data, as exemplified by responses of Neuron 2 (**Fig. 2.8B**). The resulting model was somewhat able to account for the compressed response range for outlines (compare neuronal data in **Fig. 2.4B** to model in **Fig. 2.8B**), but a strong positive correlation (r=0.69, p<0.001) between model responses to filled and outline shapes persisted that was not

present in typical data. The same held for Neuron 3: the model could account for the compressed range of responses to filled shapes (compare neuronal data in **Fig. 2.4C** to model in **Fig. 2.8C**), but a significant correlation (r=0.71, p<0.001) remained between filled and outline responses.

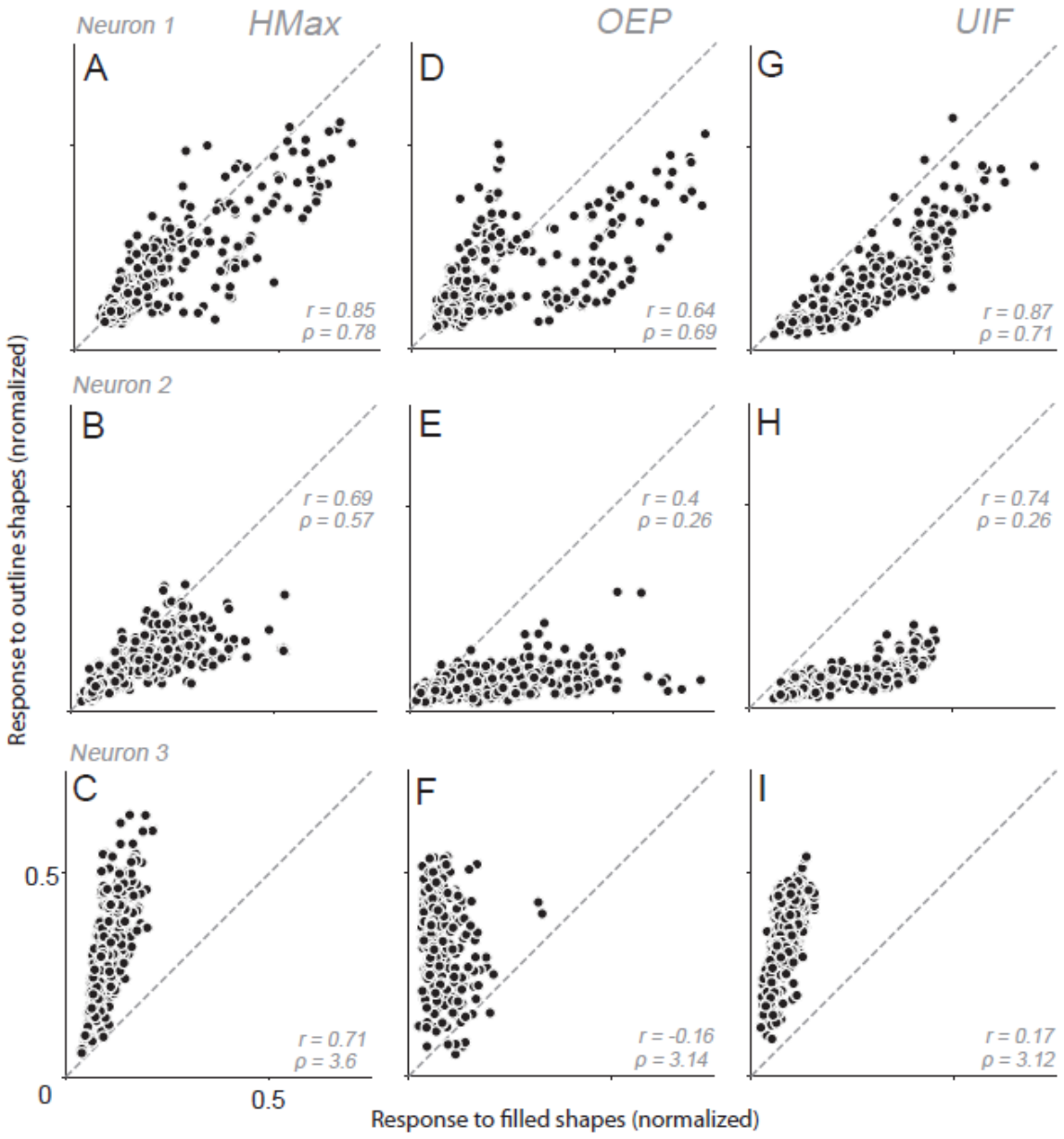


**Figure 2.8**. *Model comparison for example neurons.*

(A-C) Predicted responses to filled and outline stimuli for HMax model trained on responses to both stimulus types (*cf.* **Fig. 2.4** A-C, for corresponding scatter plots based on observed responses). For all three example

neurons, the best HMax model (even-symmetric S1 filters for A and C, odd-symmetric for B) predicted responses that were reasonably well-correlated with observed data: (A) $r_{fit}$ = 0.76; (B) $r_{fit}$ = 0.57; (C) $r_{fit}$ = 0.80. But they failed to capture the poor correlation between filled and outline responses for Neurons 2 and 3 (compare r values listed in the panels with those in **Fig. 2.4B-C**). The models also failed to capture the compressed range of outline responses for Neurons 1 and 2 (compare $\rho$ values listed in the panels with **Fig. 2.4B-C**). Models included templates with: (A) 23 C1 units; (B) 19 C1 units; (C) 21 C1 units.

(D-F) Predicted responses based on the oriented edge polarity (OEP) model. Predicted responses were well-correlated with observed responses: (D) $r_{fit}$ = 0.76; (E) $r_{fit}$ = 0.79; (F) $r_{fit}$ = 0.83. These values, as well as the fill/outline response correlations, and relative response ranges (listed in the panels) were comparable to the values based on observed data (see **Fig. 2.4A-C**). OEP models presented here were based on: (D) even-symmetric S1 filters and a template with 24 C1 units; (E) odd-symmetric S1 filters and a template with 24 C1 units; (F) odd-symmetric S1 filters and a template with 23 C1 units.

(G-I) Predicted responses based on unoriented interior fill (UIF) model. Here again, predicted and observed responses were well-correlated: (G) $r_{fit}$ = 0.74; (H) $r_{fit}$ = 0.76; (I) $r_{fit}$ = 0.84. These values, the fill/outline response correlations, and relative response ranges (listed in the panels) were comparable to the values based on observed data. UIF models presented here were based on: (G) odd-symmetric S1 filters and a template with 24 C1 units; (H) even-symmetric S1 filters and template with 25 C1 units; (I) odd-symmetric S1 filters and a template with 12 C1 units. All r values reported here are statistically significant at $p < 0.01$.

To overcome this limitation of the original model, we consider two modifications aimed at making additional information about interior fill more explicitly available to the V4-like S2 (shape template) stage of the HMax model.

In the original model, C1 units were phase insensitive (analogous to V1 complex cells) due to max-pooling across phase of the simple S1 Gabor units (**Fig. 2.2C**), resulting in a loss of information about the interior of the stimuli. To preserve phase information, we removed this operation (**Fig. 2.2D**) to create the oriented edge polarity (OEP) variant of the model, which contains oriented C1 channels sensitive to contrast polarity of edges. This model has twice as many C1 channels available as inputs to the shape-template stage (S2) during the fitting procedure, but the rest of the fitting procedure remained the same. Thus, the model was free to assign weights to either or both of the phases of a filter with a particular orientation and position in space. The resulting fits of the OEP model

for the example neurons are shown in **Figure 2.8** (**D-F**). For the fill-preferring Neuron 2, the OEP model was better able to capture the lack of correlation ($r = 0.4$, $p<0.001$) and the compressed response range for outlines (**Fig. 2.8E**; *cf*. **Fig. 2.4B**). Similarly, this phase-preserving model better replicates the behavior of the outline-preferring Neuron 3 (**Fig. 2.8F**; *cf*. **Fig. 2.4C**). On average across the population, the OEP model's ability to describe the data significantly improved compared to the original HMax model (pairwise t-test, $p < 0.001$; **Fig. 2.9A**, green points).

As an alternative to preserving phase information in the orientation tuned C1 channels, we created a variant, the unoriented interior fill (UIF) model. Unlike the first modification which incorporated information about interior fill in the phase of oriented units, in the UIF model we introduced a new set of unoriented difference-of-Gaussian S1 filters that drove a set of phase-pooling C1 channels (**Fig. 2.2E**). We reasoned that this would allow the S2 shape template to detect the interior of the object separately from the oriented boundary, giving the model an ability to discriminate filled and outline stimuli. These model predictions to filled and outline shapes are shown for Neurons 1-3 in **Figure 2.8** (**G-I**) for the best fit models. The resulting model predictions were similar to those of the OEP model. In particular, access to unoriented units allowed the model to replicate the bias of Neuron 2 for filled shapes (**Fig. 2.8H**; *cf*. **Fig. 2.4B**) and the bias of Neuron 3 for outlines (**Fig. 2.8I**; *cf.* **Fig. 2.4C**), while substantially reducing the correlation between filled and outline shapes.

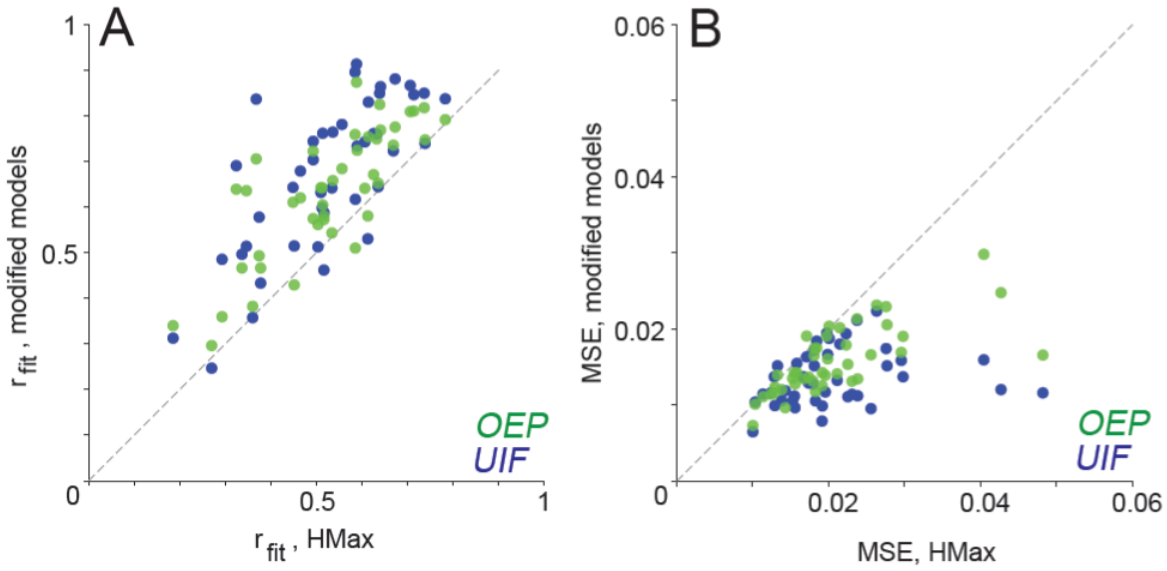


**Figure 2.9**. *Population model results.*

(A). Correlations between predicted and observed responses, $r_{fit}$, for HMax model fits (abscissa; one model per neuron, values reported are averaged across even- and odd-symmetric S1 filters) are plotted against corresponding values for OEP (green) and UIF (blue) models. Correlation values were significantly greater for OEP and UIF models compared to HMax (pairwise two-sample t-test, $p<0.001$); values for UIF were also significantly greater than for OEP model ($p = 0.0006$).

(B). Mean-squared error (MSE) associated with responses predicted by the HMax model (abscissa) are plotted against UIF and OEP models (ordinate). In most cases, OEP and UIF models were associated with smaller MSE values. Neurons that were well-described by the HMax model (MSE $<$ 0.02) showed limited reduction in MSE values with the addition of information about interior fill. The UIF model was associated with significantly lower MSE values than the HMax and OEP models (pairwise two-sample t-test, $p < 0.001$ and $p = 0.0002$).

Across the population of 43 neurons, the UIF model fit the data significantly better than the original HMax model (pairwise t-test, $p < 0.001$; **Fig. 2.9A**, blue points) and the OEP model ($p = 0.0006$; **Fig. 2.9A**, compare blue and green points). Although more units were available to the OEP and UIF models, the model comparison is supported by the constraints of the fitting process (limiting each model fit to assign weights to 2-25 C1 units)

and by the fact that fewer C1 units were assigned to each neuron by the more complex models (median = 22±6 unit weights were assigned in the HMax model fits, compared to median = 19±6 for the OEP and UIF models).

For the neurons in our population that were already well-described by the original HMax model (indicated by low MSE values in **Fig. 2.9B**, abscissa), the performance of the HMax model was indistinguishable from that of the OEP and UIF models (points near the line of equality in **Fig. 2.9B**). In contrast, for neurons whose behavior was poorly captured by the HMax (*e.g.* MSE $>$ 0.02), our modifications showed sizable improvements in performance.

In designing the OEP and UIF models, we attempted to incorporate additional information about interior fill within the HMax framework. As illustrated by the model fits to Neurons 2 and 3 above, the population-wide performance improvements of the OEP and UIF models could indeed be attributed to their better ability to predict the responses of neurons modulated by interior fill. Critically, the original HMax model was limited in its ability to capture the relative ranges of responses to filled and outline stimuli as quantified by $\rho$ (**Fig. 2.10A**). Particularly for neurons that responded poorly to outline stimuli, *i.e.* when $\rho << 1$, the $\rho$ of responses predicted by the HMax model consistently deviated from the line of equality.

In contrast, the OEP and UIF models were better able to match the relative ranges of responses for filled and outline stimuli, especially for the most fill- and outline-preferring neurons (**Fig. 2.10B**).

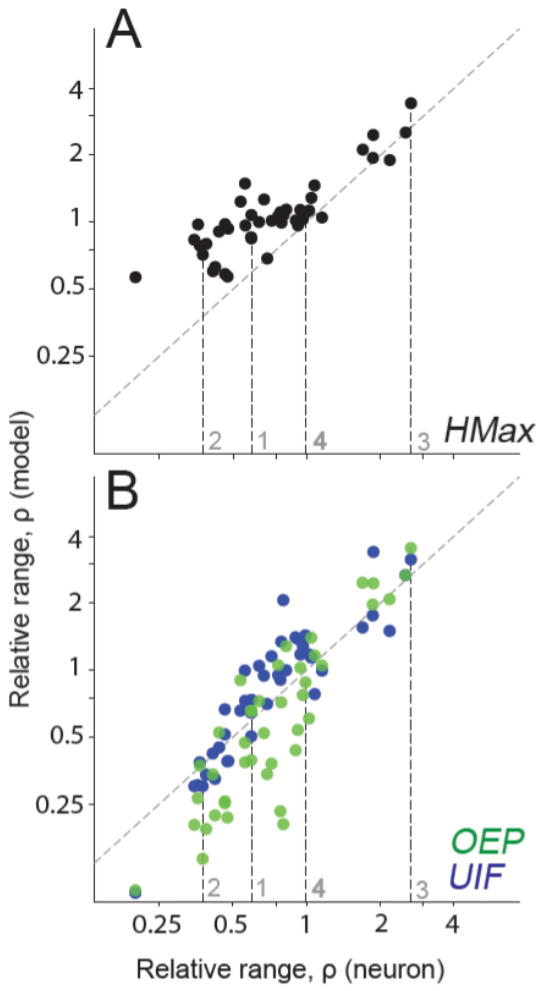


**Figure 2.10**. R*elative range of predicted versus observed responses.*

Relative range of responses to filled versus outline stimuli predicted by the HMax (A), UIF (B, blue) and OEP (B, green) are plotted against relative response range based on observed data.

When observed $\rho$ values are > 1, the HMax predictions (A) lie close to the diagonal. But this was not the case when observed $\rho$ was less than 1. $\rho$ values based on OEP and UIF (blue) models closely track the line of equality, meaning responses predicted by these models match the fill- or outline preference of many neurons.

Values reported for here are averaged across S1 phases (even- and odd-symmetric). $\rho$ values for example neurons 1-4 are indicated.

### *2.3.4 Summary of results*

We found two straightforward modifications to the HMax model that allowed it to better account for both decreased correlations between responses to filled and outline shapes and a difference in dynamic ranges in responses to these stimuli, as observed in our V4 data. Thus, in addition to statistically significant improvements in performance, these models make important qualitative improvements that are able to account for modulation by the presence or absence of surface fill.

## 2.4 DISCUSSION

### *2.4.1 Representation of interior fill in area V4*

To deepen our understanding and improve models of mid-level form processing, we investigated how a surface attribute of an object, specifically the interior fill of simple shapes, influences responses of form-selective V4 neurons. We found that very few neurons responded similarly to filled and outline stimuli. Instead, a majority of neurons responded strongly and selectively to one class of stimuli (filled or outline) but poorly to the other. V4 neurons that exhibited differential responses to filled and outline stimuli were also form-selective suggesting that the mid-level representation of form involves a complex combination of boundary and surface features.

These results demonstrate that a prevalent assumption in form encoding models, that form selectivity is built on boundary orientation signals regardless of phase, is not appropriate for a majority of neurons in area V4, where form encoding relies both on boundary conformation and surface fill. Based on these experimental findings, we propose modifications to an existing V4 HMax model that can capture the observed diversity in V4 responses.

### *2.4.2 Consistency with past experimental studies*

Tuning for form and surface properties have typically been investigated independently in V4—many past studies have demonstrated that V4 neurons are sensitive to visual form (Desimone and Schein, 1987; Kobatake and Tanaka, 1994; Gallant *et al.* 1996; Pasupathy and Connor, 1999, 2001) while many others have

demonstrated sensitivity to other object properties including luminance contrast (Bushnell *et al*., 2011b), chromatic contrast (Schein and Desimone, 1990; Conway *et al.*, 2007) and texture (Ohkazawa *et al.*, 2015; Goda *et al*., 2014). Our results document the strong influence of surface characteristics on responses of form-encoding V4 neurons. Form-selective V4 neurons lie along a continuum, ranging from neurons that respond preferentially to filled stimuli to those that prefer outlines. Consistent with past findings in IT cortex (Ito *et al*. 1994; Kayaert *et al.*, 2003), our results indicate that responses of form-selective V4 neurons are better explained when surface information is taken into account.

It is unlikely that the preference for filled stimuli at more eccentric locations arose simply because outlines may be less perceptible: preference for filled stimuli persisted even when we doubled outline widths. We also verified in early experimental sessions that animals could discriminate outline shapes presented at our typical linewidths and across typical eccentricities of our recordings.

A small subset of neurons increased their firing rates in response to both filled and outline stimuli but did not display strong form-selective responses, nor consistent modulation by interior fill. These neurons may be sensitive to a stimulus property that was not directly assessed by our experimental design, and perform a function unrelated to the representation of object boundary or interior fill.

### *2.4.3 Building models of the ventral stream*

Experimental and computational work over the past several decades has yielded numerous models for how signals from the LGN might be transformed into the simple and complex cell response properties observed in V1. However, it has been more difficult to

capture the computational architecture of the transformation from V1 to V4, and past biologically-plausible models have relied simply on boundary form information to predict V4 responses (Cadieu *et al*., 2007; Rodriguez-Sanchez and Tsotsos, 2012). Our experimental results provide a key insight: that explicit surface contrast and form information is necessary to build accurate models of V4 responses.

Because filled and outline stimuli differ in spatial frequency (SF) content, modulation of form responses could theoretically be implemented by access to SF information. For example, the spectral receptive field model (David *et al*., 2006) model could differentiate between filled and outline stimuli, but it cannot capture V4-like tuning for curvature (Oleskiw *et al.*, 2014). The inferior performance of the original HMax model suggests that joint tuning for shape and preference for fill/outline cannot be captured on the basis of differences in SF. However, the units defined in the HMax model cover a limited range of spatial frequencies – less than 1 octave (from 1.0 to 1.9 cyc/deg). Adding more low-SF oriented units to the HMax model to expand this range produces results that are similar to the OEP and UIF models in their ability to match observed responses. This limitation of the HMax helps explain the success of the UIF model, since unoriented units provide a source of low-SF information.

### *2.4.4 Improvements to the HMax model*

To provide the model with direct access to surface fill information, we tried two strategies: preserving phase information in the C1 units of the model, and including units with unoriented spatially opponent (center-surround) RFs in the S1 layer. By including additional information about surface characteristics, these models are able to account for

selectivity to object fill without compromising the ability to capture shape-selective responses. Each of these approaches is a simple modification to the HMax model, but by jointly representing surface and form information in the OEP and UIF models, we were able to achieve better model fits without increasing the number of C1 units driving the shape-selective templates (S2 units).

While the UIF and OEP models were equally successful at capturing the diversity in our V4 data with regard to responses to filled and outline stimuli, the two models exhibit complementary strengths in their ability to explain other V4 response properties.

First, previous neurophysiological studies suggest that V4 neurons encode parts of a shape boundary in an object-centered reference frame (Pasupathy and Connor, 2001). Demonstrations of position and size-invariant tuning for contour segments in V4 (El-Shamayleh and Pasupathy, 2016) are also consistent with representations in an object-centered reference frame. Although the original HMax model seeks to reflect the shape selectivity and translation invariance in responses of V4 neurons, it does not represent boundary curvature with respect to the object center; *e.g.* the contour templates of HMax model units can be driven by the same boundary regardless of whether it belongs to the foreground (object) or the background. By adding unoriented filters to the contour templates of the original HMax model, the UIF model could achieve object-centered encoding. Specifically, if the unoriented filters reflect the interior fill of the stimulus, then the position of the contour template relative to the position of the unoriented filters could represent object-centered encoding.

Second, many V4 neurons exhibit preference for the luminance contrast of a filled shape—some neurons respond best to stimuli brighter than the background, others to

stimuli darker than the background and still others to stimuli equiluminant with the background (Bushnell *et al.*, 2011b). Because of the phase-pooling operation at the C1-level, the original HMax model cannot capture this diversity in luminance contrast preference; since both oriented and unoriented units in the UIF model are also phase-pooled, this model modification also fails to capture this property of V4 responses**.** In contrast, our OEP model, where phase information is preserved at the C1 level, could successfully reproduce V4 neuronal preference for different luminance contrasts.

One critical difference between the OEP and UIF models is the relationship between the representation of contour and surface characteristics—in the former, they are tied together while they are dissociated in the latter. Thus the UIF model is more consistent with illusions that suggest a version of scene perception where surface characteristics are filled into regions defined by boundaries (Grossberg and Mingolla, 1985; Pinna and Grossberg, 2005). Given the complementarity in our UIF and OEP models in terms of their ability to capture V4 response properties, it is conceivable that both strategies of preserving information about interior fill may be instantiated across the V4 population.

Both of these modifications also appear to be biologically-plausible. While an ideal model of a complex cell is phase insensitive, most studies suggest that complex cells in V1 and V2 show a range of sensitivities to phase (Levitt *et al.*, 1994; Mechler and Ringach, 2002).

*2.4.5 Representation of interior fill in area V4 reflects perceptual demands*

We studied the responses to outline stimuli not only because we wanted to build more accurate models of V4 neurons, but also because outline stimuli are an important, perceptually relevant stimulus class (Hochberg and Brooks, 1962; Allen *et al.,* 2010). The human visual system is remarkably adept at object recognition based on outlines: even young children can recognize line drawings devoid of surface fill. Outline objects must be recognized solely on the basis of their boundary shape, and boundary-selective responses in area V4 are a natural candidate to underlie this behavior. However, our results suggest that a large fraction of neurons in V4 is poorly driven by outline stimuli, which at first appears to conflict with the human facility at recognizing outlines. Importantly, most V4 neurons that responded well to outlines were within 4° of the fovea – the region of the visual field likely to be most relevant for reading and recognition (Latham and Whitaker, 1996; Rosenholtz, 2016). Moreover, while humans are good at recognizing isolated objects from outlines, clutter makes outline based recognition hard; in such cases surface characteristics are critical for recognition. Thus, the preponderance of fill-preferring cells may reflect the importance of surface characteristics in interpreting visual scenes with clutter.

In natural vision, the information that reaches our eyes has to be parsed into a meaningful arrangement of regions and objects. This process, called image segmentation, is one of the most challenging computations accomplished by the visual system. Computer vision and psychophysical studies suggest that segmentation is most effective when both contour- and region-based properties are considered (Leung and Malik, 1998; Mumford *et al.*, 1987; Mel, 1997; von der Heydt and Pierson, 2006; Hansen

and Gegenfurtner 2007). Thus, the joint encoding of surface and boundary characteristics in a majority of V4 neurons may reflect this algorithmic preference. Thus, these V4 neurons may play a role not just in form encoding but also in scene parsing.

# Chapter 3. Aim 2: Influence of behavioral engagement on feature selectivity in area V4

## 3.1 INTRODUCTION

### *3.1.1 Feature selectivity and influence of cognitive factors*

As discussed in the previous chapter, responses of neurons in primate visual area V4 reflect different aspects of visual objects, including form, color, brightness, and texture. Tuning for these object properties has typically been investigated in studies where the animal fixates a central location on the screen, and stimuli are presented in the receptive field (RF) of the neuron. Such characterizations have formed the basis of much of our knowledge about the encoding of visual stimuli throughout the ventral pathway of visual processing, which comprises brain areas involved in object recognition. However, as information about stimuli progresses along the pathway from the primary visual cortex to inferotemporal cortices, neuronal responses display an increasingly prominent influence of cognitive factors (Buffalo *et al.*, 2010), which is thought to enable the animal to perform visually-guided tasks.

It is largely unknown whether object representation, as we understand it from fixation studies, is similar when animals are engaged in a behavioral task - that is, how does selectivity for object features, *e.g.* form and color, compare in these two contexts?

### *3.1.2 Insight from studies of attention effects in area V4*

Many previous studies investigated the influence of cognitive factors in area V4 in the context of attentional paradigms (*e.g.* Haenny *et al*. 1986; McAdams and Maunsell

1999, 2000). Such studies generally compared responses to stimuli within or outside the neuronal RF as the animal was cued to attend to a spatial location and/or an object feature (*e.g.* detect the blue object or the dots moving upward). In these contexts, V4 neurons show enhanced responses to the attended stimulus location or feature; in studies where response magnitude was examined alongside feature selectivity, tuning during the "attended" condition could be explained by applying a single multiplicative factor to the responses during the "unattended" condition (*e.g.* McAdams and Maunsell, 1999). However, the animal typically performed the same behavioral task in both the attended and unattended conditions; the nomenclature refers to whether the cued spatial location overlaps the neuron's RF. Studies of V4 responses during visual search tasks have reported a similar influence of behavioral goals on responses of V4 neurons, which are influenced by the target feature (*e.g.* Ogawa and Komatsu, 2004). However, studies using visual search task paradigms also typically compare neuronal responses between two conditions where the animal is performing the same behavioral task.

### *3.1.3. Potential influence of task context on feature selectivity*

It is possible that when an animal views the same stimuli while passively fixating, neuronal responses could be similar to the responses in the unattended condition in the paradigms described above; that is, responses during behavior are scaled versions of responses during fixation and feature selectivity is maintained. Alternatively, behavioral engagement may induce a change in feature selectivity, including a shift in the tuning peak (indicating a change in the neuron's preferred feature) or changes in the tuning width (indicating an increase in discriminability between preferred and non-preferred features if

tuning narrows, or the converse if it broadens). Such changes have been observed in area V4 in the context of an easy *vs*. difficult discrimination task (Spitzer *et al.*, 1988) and during feature-based attention (David *et al.*, 2008), as well as in the dorsal stream of visual processing, where area LIP (lateral intraparietal area) neurons display shifts in selectivity for stimulus color and direction of motion, depending on the feature relevant for the task (Ibos and Freedman, 2014).

### *3.1.4 Chapter overview*

In the following chapter, I will examine how the shape and color selectivity of V4 neurons is affected by task context. An ideal study would compare neuronal responses to stimuli during a fixation task, where the animal was simply fixating a central location on the screen, and during two discrimination tasks: one where animals are asked to discriminate shape, and another where animals are asked to discriminate color. Experimental constraints exist both in training the animal to robustly judge identical stimuli with changing discriminanda, and in the time required to obtain enough data for a comparison across 3 tasks while maintaining the quality of isolation of single neuron responses. Thus, I will present a more restricted comparison of responses of V4 neurons to the same visual stimuli presented while animals judged stimulus shape in a sequential discrimination task, and while they simply fixated.

I will examine whether neuronal responses during the behavior task exhibited changes in response that were dependent on the firing rate during the discrimination task and retained feature selectivity, or were dependent on stimulus features and reflected a

change in tuning for stimulus form and color. Finally, I will discuss whether these changes related to the behavioral relevance of the feature to the discrimination. In our discrimination task, stimuli differed in both shape and color, but only stimulus shape was relevant for accurate performance in the task; for several neurons I collected additional responses during a task where color was relevant, and shape irrelevant, for accurate performance.

## 3.2 METHODS

### *3.2.1. Surgical methods*

We implanted 2 adult male macaque monkeys (*Macaca mulatta*) with custom headposts and chambers positioned over dorsal area V4 (left hemisphere). The placement of the chamber over the prelunate gyrus and subsequent craniotomy were guided by structural MRI (for details, see Bushnell *et al.,* 2011a). All animal procedures conformed to the National Institutes of Health guidelines and were approved by the Institutional Animal Care and Use Committee at the University of Washington.

### *3.2.2. Visual stimulus presentation*

Stimuli were presented against a uniform gray background (luminance 5.4 cd/$m^2$) using a spectrally calibrated (PR650, PhotoResearch) CRT monitor positioned 45.4 or 56 cm away (subjects 1 and 2, respectively). The animal fixated a small white dot in the center of the screen within a window of radius 0.75°-1°. Eye position was monitored using an infrared eye tracking system (EyeLink 1000; SR Research) and coordinated with stimulus presentation using custom software based on Pype (developed by James Mazer).

*3.2.3. Task design*

*3.2.3.1 Fixation task.* Each trial of the fixation task (see schematic in **Fig. 3.1A**) began with the presentation of a fixation spot. Once the animal acquired fixation, a sequence of 4-5 stimuli were presented with a 200 ms blank interval preceding each stimulus. For the first 12 neurons we recorded, each stimulus was presented for 300 ms**;** this was later changed to 250 ms to better match the timing of stimuli in the discrimination task (see below). A total of nine stimuli (3 shapes x 3 colors) customized to the shape and color preferences of the neuron under study (see Stimulus Selection, below) were used to probe responses during the fixation task. Stimuli were presented in random order, and each was shown 20 times within a block of fixation task trials**.** The animal was rewarded with drops of juice for successful fixation for the entire duration of a trial (typically 2000-2500ms).

*3.2.3.2 Discrimination task.* Animals were also trained on a sequential shape discrimination task (**Fig. 3.1B**). Each trial began with the presentation of a fixation spot. Once the animal acquired fixation, a "reference" stimulus was presented at central fixation for a duration of 600 ms. After an inter-stimulus interval of 200 ms, a "test" stimulus appeared within the receptive field (RF) of the V4 neuron. Simultaneously, two small target dots appeared to the left and right of the fixation spot, each 6 degrees of visual angle away from fixation. The animal then had up to 1500 ms to report whether the shape of the test stimulus was the same or different from the reference stimulus via a saccadic eye-movement to the right or left target dots, respectively. The test stimulus disappeared as soon as the animal's eye left the fixation window; test stimulus duration varied from trial to trial, dictated by the reaction time of the animal. Mean reaction times, measured

as the time at which the animal left the fixation window, were 260±23 ms across both animals. The same 9 stimuli used during the fixation task served as the reference and test stimuli in the discrimination task. Each of the 9 stimuli appeared randomly as a test stimulus for a total of 15-30 repeated presentations in each block of discrimination trials. The animal was rewarded with drops of juice for a saccade to the correct target location; the number of trials requiring a leftward or rightward saccade were balanced. Chance performance was 50%; mean animal performance was 83±5%. In our analyses below, we include only responses to stimuli where the outcome of the trial was a correctly directed saccade.

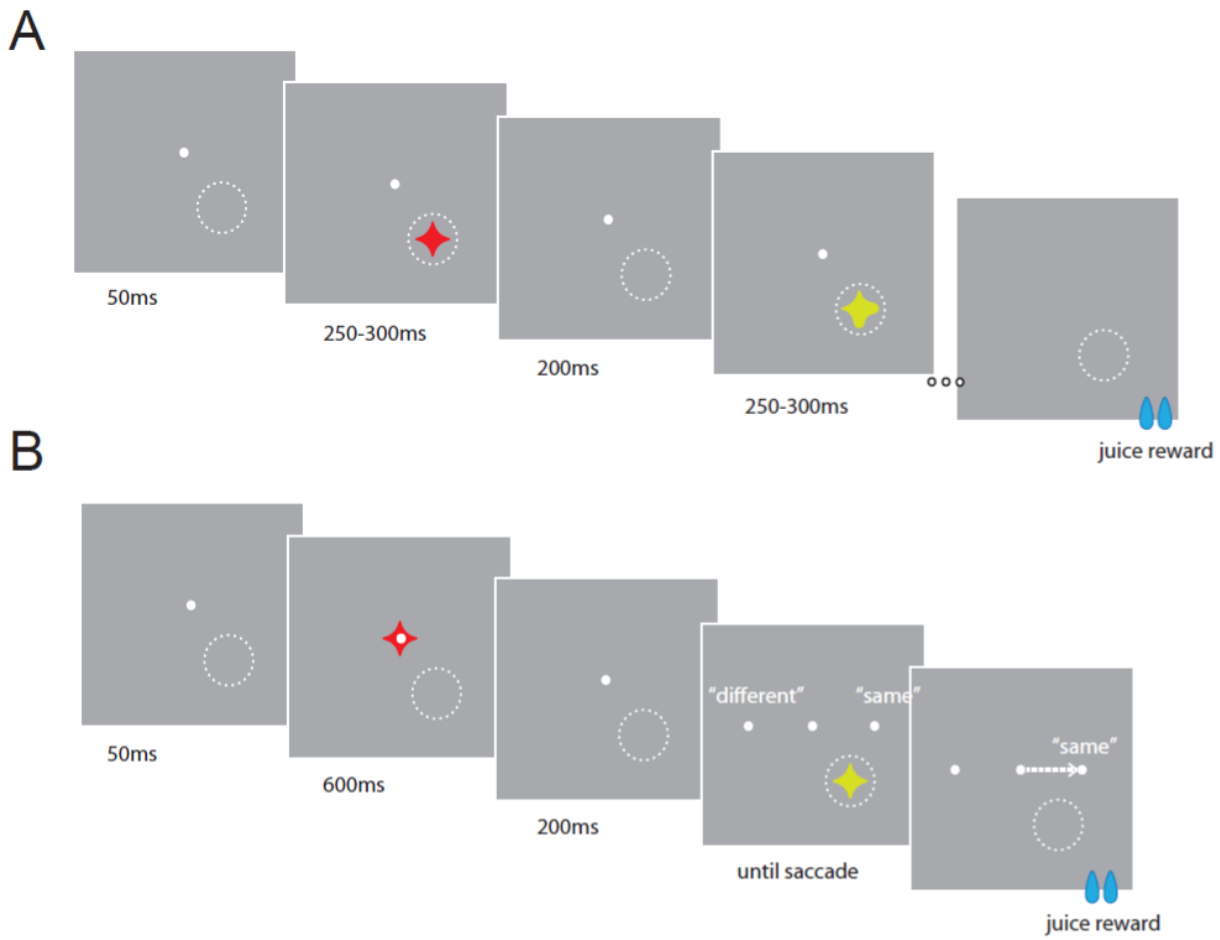


**Figure 3.1**. *Task conditions and stimulus design.*

(A). Fixation task. The white dotted circle represents the spatial extent of the receptive field (RF) of the V4 neuron under study and was not presented on the screen. The animal fixated a small dot at the center of the screen (white) while stimuli were presented in the V4 neuron's RF one at a time and separated by 200 ms. The animal was rewarded with drops of juice for continuously fixating a sequence containing 3-5 stimuli.

(B). Delayed match to sample (discrimination) task. The white dotted circle represents the spatial extent of the V4 neuron under study and was not presented on the screen. The animal fixated a small for at the center of the screen (white) and a reference stimulus was first presented in the same location as the fixation spot for 600 ms, followed by a 200 ms interval. A test stimulus was then presented in the V4 neuron's RF along with two saccade choice target dots; the animal then had 1.5s to report whether the shape of the test stimulus matched (saccade to the target on the right, "same") or did not match the reference stimulus (saccade to the target on the left, '"different") and was rewarded with drops of juice for a correct saccade. In the example trial

shown here, the two stimuli had the same shape so the animal would be rewarded for a saccade to the rightward target, represented by arrow. On average, animals took 260 ms to report their choice.

(C). Example stimuli from one session. We selected 3 shapes for each neuron based on preliminary screening, with a shape that elicited a robust response from the neuron; a shape that elicited a poor response from the neuron; and a shape that elicited an intermediate level of response. We used the same process to choose 3 colors (differing in hue and luminance), and created a set of 9 stimuli with unique shape-color pairings that were then presented in the task contexts schematized in (A) and (B). Ideally, the resulting set of stimuli was such that from the most preferred stimulus (top left corner) to the least preferred stimulus (bottom right corner), the neuron's spiking responses spanned much of its dynamic range. Colors shown are adjusted for visibility; actual colors were similar and calibrated for the specific monitor used.

### *3.2.4 Data collection*

Each day, we lowered a single tungsten microelectrode (FHC) using a stepper motor drive (Gray Matter Research). We amplified and filtered the signals from the electrode and sorted waveforms (Plexon Systems) to identify single units. Here, we include responses from 83 well-isolated neurons for which RF eccentricities ranged from 1.6° to 7.3°.

*Stimulus selection.* For each neuron in our dataset, we first assessed the spatial extent of the RF manually using a variety of shapes and colors under mouse control. We then characterized each neuron's color/luminance preference either manually (6 neurons across 4 sessions) or via an automated protocol that measured neuronal responses to a single shape presented in 25 colors that uniformly sampled the CIE space, and at 3-4 luminance levels (see Bushnell *et al*., 2011b). To assess shape selectivity, we presented 25 2D shapes (a subset of the shapes used by Pasupathy and Connor, 2001) at 8 rotations, for a total of 200 shape stimuli.

Based on the color/luminance and shape preferences of each neuron, we chose three shapes and three colors that represented the preferred, intermediate and non-preferred values along each of the two feature dimensions for each neuron. We then created a set of 9 stimuli by combining the shapes and colors thus selected for use during the fixation and behavioral tasks (see Figure **3.1C** for an example set from one session). When responses of more than one neuron were recorded simultaneously (14 sessions in animal 2), stimuli were tailored according to the preferences of only one of the neurons.

Position of the test stimuli during behavior and all stimuli during fixation were jittered by up to 5 pixels about the center of the neuron's RF. Stimuli were scaled such that all parts of the stimuli were within 80% of the estimated RF diameter based on data in Gattass *et al.* (1988); reference stimuli in the discrimination task were size-matched to the RF-based scaling of the test stimuli.

For each neuron, fixation and behavioral tasks were conducted in alternating blocks starting with the fixation task. Each fixation block was 45-60 trials long, and each discrimination block was 144-288 trials long; as described above, 3-4 stimuli were presented per fixation trial, while one stimulus was presented in the RF per behavior trial. The characteristics of the fixation spot identified the task: a white cross indicated a fixation task while a blue square indicated a discrimination task. Here we include responses from neurons where we successfully captured at least 2 blocks of both tasks (median: 3 blocks for each task).

*3.2.5 Data Analysis*

*3.2.5.1 Computing neuronal responses.* Stimulus onset and offset times were detected with a photodiode. To construct peristimulus time histograms for each neuron (PSTH, as in **Fig. 3.2A**), responses in all blocks of one task type (fixation or discrimination) were aligned to stimulus onset and 1-ms bin responses were smoothed with a Gaussian kernel ($\sigma$ = 10 ms). To compute responses across time for each stimulus, we computed the mean spiking rate on each trial in the time window of 50-200 ms from stimulus onset. This time window was chosen to coincide, at the earlier end, with mean latency of signal arrival in V4 (Schmolesky *et al*., 1998; Zamarashkina *et al.*, 2017). The later bound of 200 ms was chosen to account for the influence of reward anticipation (in both tasks) or preparation for saccadic eye movements (in the discrimination task, where time to saccade initiation was 260±23 ms). To construct tuning schematics (as in **Fig. 3.2B**), we further averaged responses across multiple presentations of each stimulus.

To obtain feature-specific average PSTHs (*i.e.* shape- or color-specific responses) we first averaged across all stimuli with a particular feature, *e.g.* the preferred shape in all colors. Each feature-specific PSTH thus contained 3 responses: one for the preferred feature, one for the intermediate, and one for the nonpreferred. We normalized these PSTH traces separately for each task and neuron, and averaged the resulting PSTHs across neurons. Finally, we obtained tuning curves, as above, by computing average responses in the 50-200 ms window for each of the 3 stimuli, and normalized them such that the response to the preferred feature was 1.

*3.2.5.2 Regression models.* To compare neuronal responses to the 9 stimuli during fixation and discrimination, we used individual trial responses to fit four regression

models. Based on the body of literature examining attentional influences in V4, we expected to see a difference in responses during the discrimination task, and chose a variety of models to compare our observations to previously published results.

First, we considered the possibility that neuronal responses during the behavioral task scaled linearly relative to the responses during the fixation task, as reported in attentional studies; such a response change could be described using the linear-gain model:

$$R_{beh(s,c)} = k_1 * R_{fix(s,c)} + k_0$$

*Equation 3.1*

where $R_{beh}$ and $R_{fix}$ denote the neuronal response during the discrimination and fixation tasks respectively, and $k_1$ and $k_0$ are the slope and constant for the regression.

To perform the regression, we needed to pair trials from fixation and discrimination blocks ($R_{fix}$ and $R_{beh}$). Fewer trials were shown in the discrimination condition; thus, we first resampled responses during the discrimination task using a bootstrap procedure, randomly drawing responses from a uniform distribution (with replacement) to match the number of trials during the fixation task. We then randomly paired responses during the fixation task with responses during the discrimination task for each of the 9 stimuli ($R_{fix(s,c)}$ and $R_{beh(s,c)}$). To reduce the introduction of bias, we repeated the resampling and pairing procedure to obtain 100 sets of regression data, and performed subsequent model fitting and analyses of model performance across these sets.

Additionally, we considered a model where neuronal responses during the discrimination task also depend on the responses during the fixation task, but this relationship is non-linear (polynomial-gain model):

$$R_{beh(s,c)} = k_2 * R_{fix(s,c)}^2 + k_1 * R_{fix(s,c)} + k_0$$

*Equation 3.2*

where $k_2$ and $k_1$ are the 2nd and 1st degree regression coefficients, and $k_0$ is the regression constant. We also considered exponential and logarithmic forms of this nonlinear relationship, but did not find these models to fit our data better than the 2nd degree polynomial form.

As an alternative to firing-rate based models, we considered a case where neuronal responses during the discrimination task depend not only on responses during the fixation task, but also on the features of the stimulus in its RF. We defined feature-gain models for shape and color of the stimulus, as follows:

$$R_{beh(s,c)} = k(s) * R_{fix(s,c)} + k_0$$

*Equation 3.3a, shape-dependent feature gain*

$$R_{beh(s,c)} = k(c) * R_{fix(s,c)} + k_0$$

*Equation 3.3b, color-dependent feature gain*

Where k(s) and k(c) represent separate coefficients for each of 3 stimulus shapes or colors, respectively, and $k_0$ is the regression constant.

*3.2.5.3 Model comparison*. To compare the models above, which have different numbers of parameter constraints, we performed cross-validation using 4/5 of the neuronal data to fit each regression model, and the remaining 1/5 to test the model fit. Responses to each of the 9 stimuli were included in both the model fitting and testing sets. As described above in *Regression models*, we repeatedly randomized the pairing of responses during fixation and discrimination tasks; we performed this cross-validation procedure for each random pairing.

To quantify performance, we calculated root mean squared error (RMSE) for the test sets, and compared the RMSE distributions and average RMSE values for each model to determine the best-fitting model for each neuron. A model was determined to fit better if its average RMSE value was lowest and the distribution of RMSE values was significantly different than other models (unpaired t-test, $\alpha=0.01$; computed for pairs of models).

To obtain predicted responses during discrimination using each regression model, we first averaged model coefficients over 100 randomized pairings of fixation-discrimination trials. We then used the averaged coefficients together with the average responses to 9 unique stimuli in the fixation condition to generate predicted discrimination responses from Eqns. 3.1-4.

*3.2.5.4 Feature selectivity.* For each neuron, we estimated the strength of feature selectivity for shape and color by calculating the partitioned variance due to each feature from a 2-way ANOVA model with shape, color, and interaction terms. The relative feature selectivity was calculated as the partitioned sum of squares for each factor (shape or

color) divided by the total sum of squares. The interaction term was small for all neurons; on average, it accounted for 4% of the relative variance.

To compare the relative strength of feature selectivity, we computed a ratio of the shape- and color-selectivity as determined above. This metric is >1 for more shape- than color-selective neurons, and <1 for more color- than shape-selective neurons.

*3.2.5.5 Contribution of other task-dependent variables.* To determine whether the identity of the reference stimulus contributed to the effects we observed, we first separated neuronal responses during the discrimination task into two sets according to whether the reference stimulus matched or did not match the neuronal preference for shape and color. The first set of data contained responses to test stimuli in trials where the reference stimulus was of the preferred shape and color (reference-preferred), and the second set contained responses to test stimuli in trials where the reference stimulus had any other combination of shape and color (reference-other).

We then examined the magnitude of the reference stimulus-dependent effect, captured in the difference between responses in the reference-preferred and reference-other datasets, relative to the overall responsiveness of the neuron using the following ratio:

$$\frac{\sum\left|Resp_{ref-pref} - Resp_{ref-other}\right|}{\sum Resp_{beh}}$$

*Equation 3.4a*

where $Resp_{ref\text{-}pref}$ is the mean response for each of 9 test stimuli for the subset of data where the reference stimulus had the preferred shape and color, $Resp_{ref\text{-}other}$ is the mean

response to each of the 9 test stimuli for the subset of data where the reference stimulus had other shape and color combinations, and $Resp_{beh}$ is the average response to each of the 9 stimuli during the discrimination trials. For a neuron with a mean response of 50 spk/s, a ratio with a value of ~1 represents a mean stimulus response difference of 5 spk/s between the two data subsets.

We performed the same analysis to determine whether the motor plan being prepared by the animal in the behavioral condition (a saccadic eye movement) contributed to the effects we observed. We separated neuronal responses during the discrimination task into two sets according to the saccade outcome of the task, and compared the magnitude of the saccade-related effect using an equivalent ratio to Eqn. 3.4a:

$$\frac{\sum\left|Resp_{right} - Resp_{left}\right|}{\sum Resp_{beh}}$$

*Equation 3.4b*

where $Resp_{right}$ is the mean response for each of 9 test stimuli for the subset of data where the trial ended in a rightward saccade, $Resp_{left}$ is the mean response to each of the 9 test stimuli for the subset of data where the trial ended in a leftward saccade, and $Resp_{beh}$ is the average response to each of the 9 stimuli during the discrimination trials. For a neuron with a mean response of 50 spk/s, a ratio with a value of ~1 represents a mean stimulus response difference of 5 spk/s between the two data subsets.

## 3.3 RESULTS

To understand how behavioral engagement influences shape and color tuning in area V4, we studied the responses of well-isolated V4 neurons (n = 83; 31 from animal 1, 52 from animal 2) to stimuli presented either during passive fixation, or during behavioral engagement in a shape discrimination task. During the fixation task (**Fig. 3.1A**), a sequence of stimuli was presented within the RF of the V4 neuron while the animal maintained fixation at a central location on the screen. During the discrimination task (**Fig. 3.1B**), the animal had to judge whether two sequential stimuli had the same or different shape and reported the decision with a rightward or leftward saccadic eye movement, respectively. We compared neuronal responses to 9 stimuli in these two tasks, alternating task blocks; the stimuli, composed from combinations of three 2-D shapes and three different colors, were chosen based on preliminary characterizations to span the dynamic range of each neuron's firing rate (see Methods for details). Sample stimuli from one session are shown in Figure 3.1C.

### *3.3.1 Responses of example V4 neurons*

Many neurons in our population displayed a change in firing rate when comparing responses during the fixation task to those during the sequential discrimination task. Below, we include three examples of neurons that demonstrate the ways in which responses changed during the discrimination task.

**Figure 3.2** shows the responses of neuron #1 (p150514) to the presentation of the same 9 stimuli while the animal performed the fixation and discrimination tasks. Figure **3.2A** displays the firing rate of this neuron as a peristimulus time histogram (PSTH; ordinate: mean FR +-SEM) aligned to the onset of the stimulus (abscissa); panels correspond to the stimuli shown in panel insets. Average responses to the same stimulus in all fixation task blocks are shown as black traces, and in discrimination task blocks as red traces.

When the animal was performing the fixation task, neuron #1 displayed responses that were modulated both by stimulus shape (*cf.* PSTHs for stimuli in top row) and color (*cf*. PSTHs for stimuli in left column). The most preferred stimulus is depicted in the top left panel. During the discrimination task, responses of neuron #1 displayed a suppression of firing rate (*cf.* red PSTH traces to black PSTH traces). While the animal was discriminating shape, neuronal responses retained shape and color selectivity, and neuron #1 continued to respond best to the stimulus in the top left panel.

Unlike stimuli varying in a single feature dimension such as orientation or contrast, our shape-color combination stimuli do not lie on a naturally ordered axis. To visualize the tuning curve for neurons under study, we created a tuning axis by ordering stimuli based on the magnitude of responses during the fixation task, which we averaged in the 50-200 ms window after stimulus onset (see gray regions in **Fig. 3.2A**). Figure 3.2B shows the tuning curve for neuron #1; black traces represent average neuronal responses during the fixation task for stimuli depicted on the abscissa, and red traces represent average responses during the discrimination task. The resulting tuning curve shows that the responses to stimuli during the discrimination task were lower than during the fixation

task, but ranked stimulus preference was similar. Such a change in responses may be reflect the simple scaling of responses as shown in studies of attention (*e.g*. McAdams and Maunsell, 1999).

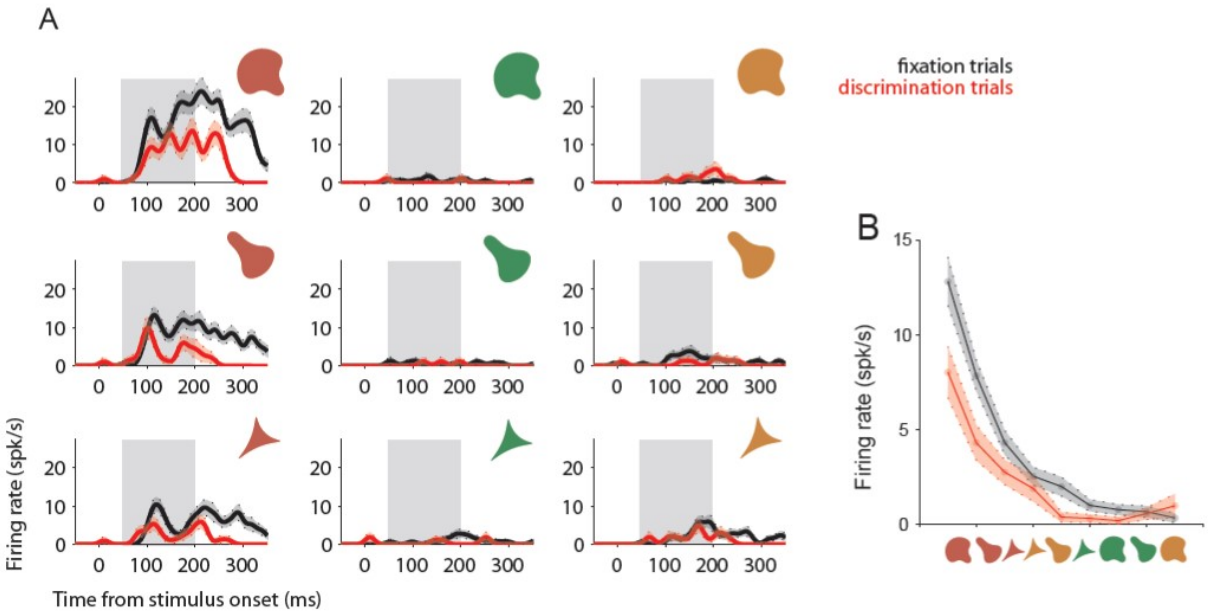


**Figure 3.2**. *Responses of example neuron #1 to stimuli in fixation and discrimination task conditions.*

(A). Peristimulus time histograms (PSTH) for responses of neuron #1 to 9 stimuli while the animal performed a fixation task (black traces) or discrimination task (red traces). The corresponding stimulus for each panel is shown in the insets; panels are arranged such that stimuli sharing the same shape are in rows, and sharing the same color are in columns. Responses of neuron #1 were both color- and shape-selective during the fixation task (compare responses to stimuli in top row; stimuli in left column). During the discrimination task, responses remained shape- and color-selective, and response magnitude decreased for all stimuli. Gray boxes indicates time window for response averaging (see below).

(B). Schematic representation of tuning for neuron #1. Each point represents the average response to the stimulus depicted along the abscissa (50-200 ms time window, see gray window in A). Abscissa is ordered according to stimulus preference, *i.e.* response magnitude during the fixation task (black trace), and responses during the discrimination task shown in red. Upper and lower bounds are SEM. Note the similar shape of the tuning schematic for the two task conditions.

In contrast, in other neurons we observed changes in responses during behavioral engagement that did not preserve the ordering of stimulus preference. Figure 3.3A shows responses of one such neuron (#2; o140605) to the 9 stimuli depicted in panel insets. When the animal performed the fixation task (black traces), this neuron displayed differential responses to both stimulus shape (*cf*. PSTHs for stimuli in top row) and color (*cf*. PSTHs for stimuli in left column), with the highest responses elicited by the stimulus in the top left panel. When the animal discriminated stimulus shape, firing rates increased for each stimulus (red traces). However, the magnitude of this change varied, although the stimulus in the top left remained the preferred one. During the discrimination task condition, stimuli with the same shape (*e.g.* panels in top row) elicited a similar magnitude of response. This result suggests that responses during discrimination are influenced more by stimulus shape than stimulus color, and thus the neuron's change in response may be feature-dependent.

This dependence is captured in the shape of the tuning curve for neuron #2, shown in Figure 3.3B. Specifically, stimuli with a certain shape elicited a larger increase in FR during discrimination than would be expected from the order of stimulus preference in the tuning curve of responses during fixation (*e.g.*, dashed lines highlight responses of stimuli in top row of **Fig. 3.3A**). This change in the tuning curve profile suggests that, unlike in neuron #1, responses of neuron #2 during discrimination were not consistent with a simple scaling of responses during fixation.

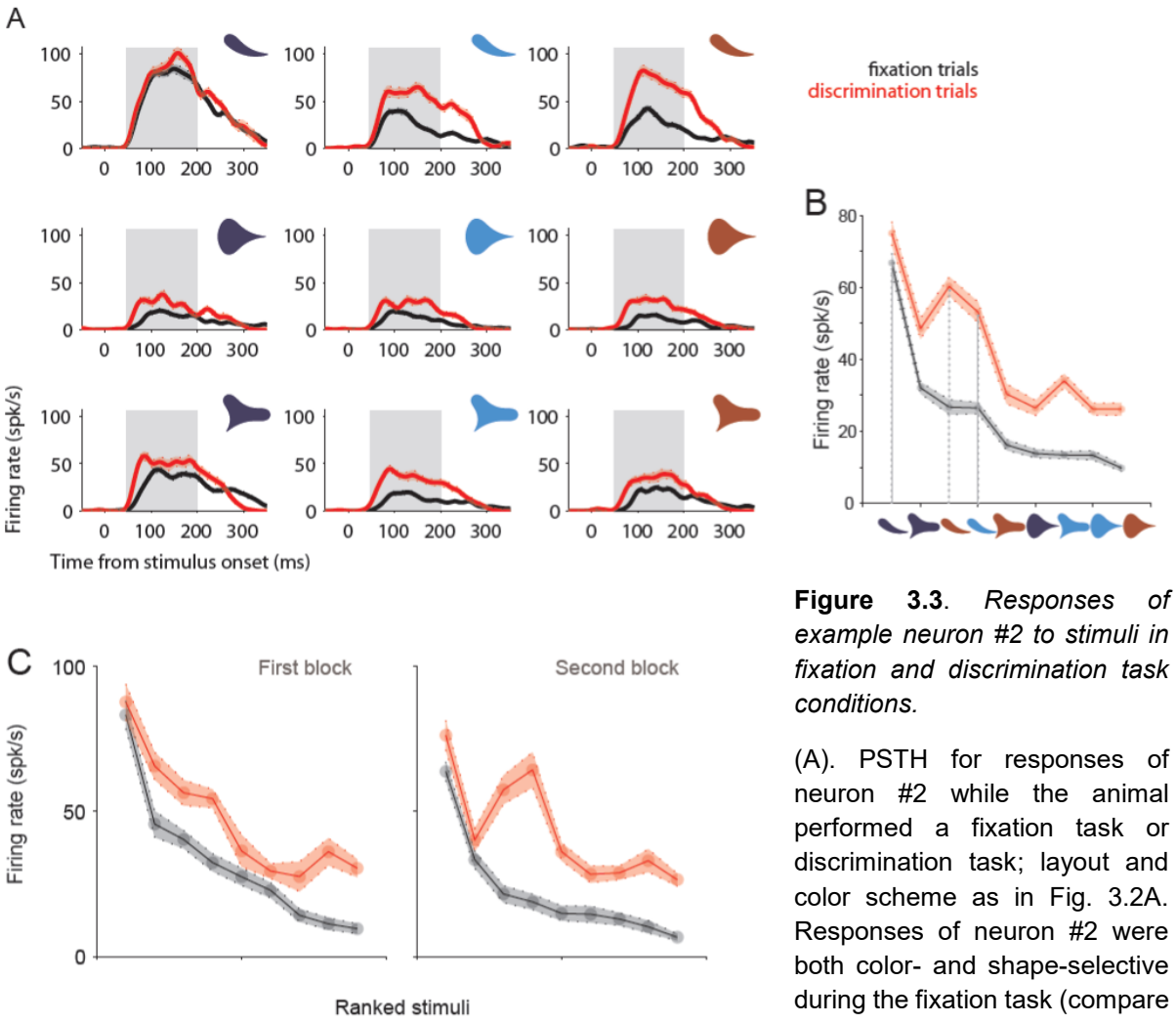


**Figure 3.3**. *Responses of example neuron #2 to stimuli in fixation and discrimination task conditions.*

(A). PSTH for responses of neuron #2 while the animal performed a fixation task or discrimination task; layout and color scheme as in Fig. 3.2A. Responses of neuron #2 were both color- and shape-selective during the fixation task (compare responses to stimuli in top row; stimuli in left column, respectively). Note that during the discrimination task, responses to all stimuli with the same shape (in rows) became more similar in magnitude across different colors. Traces are averaged across all blocks of the same task condition; upper and lower bounds represent SEM.

(B). Schematic representation of tuning for neuron #2, computed as in Fig. 3.2B. Grey dashed lines highlight responses to stimuli in top row of (A).

(C). Tuning for neuron #2 separated by block order. The tasks were presented in alternating blocks, beginning with the fixation task condition. This example is typical of stimulus tuning robustness across blocks.

This dissimilarity in the stimulus tuning during fixation and discrimination tasks was robust. The animals performed tasks in alternating blocks, and all data presented so far were averaged across at least two blocks of the same task. When examining responses of the neuron in example #2 within a particular task block, tuning during discrimination was consistently different from tuning during fixation (compare left and right panels of **Fig. 3.3C**). Robustness of response differences across task blocks was typical for all recording sessions.

Responses of an additional example neuron (#3; p150604) are shown in Figure 3.4. Like the other two examples, this neuron displayed responses during fixation (black traces, **Fig. 3.4A**) that were modulated by both shape and color (*cf.* PSTHs for stimuli in top row and left column, respectively). The stimulus with the shape and color eliciting the highest response from neuron #3 during fixation is shown in the top left panel. Similarly to the previous example neuron (#2), during the discrimination task the firing rates increased for all stimuli (red traces). However, the stimulus eliciting the highest response from neuron #3 during discrimination was no longer the stimulus in the top left panel; instead, the neuron responded best to the stimulus in the top center panel. Moreover, stimuli with the yellow color (center column) elicited a similar magnitude of response during discrimination, although the same stimuli evoked different responses during the fixation task. This result is complementary to our observations for neuron #2, suggesting that responses during discrimination were influenced more by stimulus color than stimulus shape, and thus the neuron's change in response is feature-dependent.

This dependence is reflected in the tuning curve for neuron #3, shown in Figure 3.4B. Stimuli with a yellow color elicited a larger increase in FR than expected from the

stimulus preference during fixation (*e.g.* dashed lines highlight responses to stimuli in center column of **Fig. 3.4A**). As for neuron #2, this change in the tuning curve profile suggests that responses of neuron #3 during discrimination were not consistent with a simple scaling of responses during fixation.

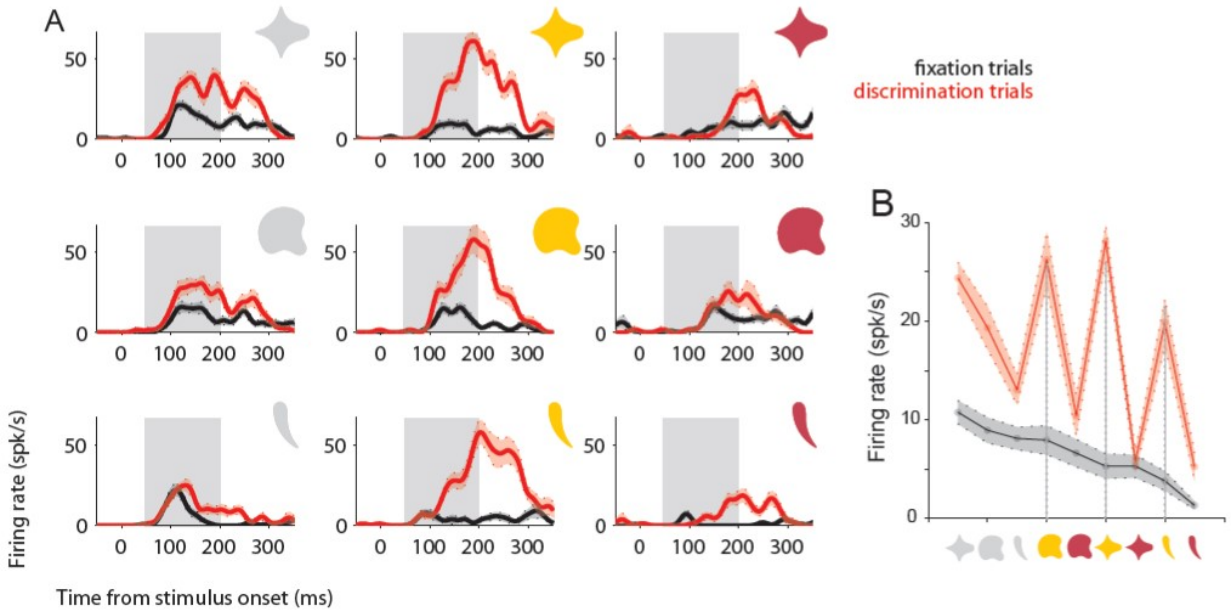


**Figure 3.4**. *Responses of example neuron #3 to stimuli in fixation and discrimination task conditions.*

Stimuli shown in light gray appeared white against the uniform gray background used in our experiment.

(A). PSTH for responses of neuron #3 while the animal performed a fixation task or discrimination task; layout and color scheme as in Fig. 3.2A. Responses of neuron #3 were both color- and shape-selective during the fixation task (compare responses to stimuli in top row; stimuli in left column, respectively). Note that during the discrimination task, responses to a less preferred color (middle column) became stronger than the originally preferred color, and more similar in magnitude across different shapes.

(B). Schematic representation of tuning for neuron #3, computed as in Fig. 3.2B. Grey dashed lines indicate responses to stimuli in middle column of (A).

*3.3.2 Modeling response scaling during behavioral engagement*

As exemplified above, we observed a variety of changes in individual neurons when comparing responses during tasks with different levels of behavioral engagement.

To quantify the changes we observed in the fixation and discrimination tasks, we performed a regression analysis using averaged responses of each neuron during each tasks to assess whether a linear, single-gain model of response modulation (Eqn. 3.1) was able to account for the responses observed in our population. Alternatively, it is possible that the influence of behavioral engagement could be explained on the basis of a nonlinear gain function that is still dependent on the firing rate during fixation. For example, the demands of the sequential discrimination task may cause a neuron to respond more strongly to its preferred than nonpreferred stimuli. To improve discriminability, responses to a neuron's preferred stimuli may be enhanced supralinearly during discrimination. Thus, we compared the predictions of a linear, single-gain regression model for each neuron to the prediction of a nonlinear, 2nd degree polynomial regression model (Eqn. 3.2).

In our population, we observed neurons such as examples #2 and #3, which displayed a change in responses that did not correlate with stimulus preference; thus, their responses during the discrimination task could not be captured by a firing-rate dependent scaling factor applied to responses during the fixation task. Therefore, we considered an alternative in which changes in responses during the discrimination task may be feature-dependent. That is, responses to stimuli with a particular shape or color may display a different magnitude of change during discrimination than responses to stimuli with another shape or color. To assess whether any neurons in our population

displayed feature-dependent responses, we compared the prediction of the two firing rate-based regression models described above to the prediction of nonlinear, shape- or color-dependent models (Eqn. 3.3a and 3.3b, respectively) which incorporate separate scaling coefficients for each stimulus shape or color.

Since the tasks we used were presented in blocks and varied in the number of trials, we resampled and randomly paired fixation and discrimination trials for each stimulus. To compare the models, which contained a different number of parameters, we performed the regressions using a cross-validation procedure. We repeated the pairing and cross-validation process 100 times for each neuron to obtain a robust distribution of performance measurements (see Methods for more detail about the fitting procedure). Below, we compare the average performance of the models over 100 iterations – first for the example neurons and then for the entire population we recorded.

We compared the observed fixation and discrimination responses of example neurons #1-3 to the predicted discrimination responses from each of the 4 models described above (**Fig. 3.5**). Discrimination responses predicted from the models are shown as open symbols and dashed lines (see Methods for details about prediction). For example neuron #1, observed responses are shown in the panels in the top row (**Fig. 3.5A-D**, solid black and red lines and symbols; same data as **Fig. 3.2B**). We used the distribution of RMSE values generated from 100 regressions (the random pairing and cross-validation procedure described above) to evaluate the performance of the regression models in pairwise comparisons. Discrimination responses of this neuron were well-captured by the linear-gain model (**Fig**. **3.5A**; mean RMSE across 100 predictions: 4.5). The polynomial-gain model and the feature-dependent models (**Figs. 3.5B-D**)

produced similar predictions (mean RMSE: 4.5-4.6), and their performance was not significantly different (unpaired t-test, $p > 0.2$ for all pairwise model comparisons; $\alpha = 0.01$). Thus, the linear-gain model provided the most parsimonious explanation of this neuron's response change due to added behavioral engagement.

Observed responses of neuron #2 are shown in the panels in the middle row (**Fig. 3.5 E-H**; solid black and red lines and symbols; same data as **Fig. 3.3B**). The prediction of the firing-rate dependent models (**Fig**. **3.5 E-F**) were similar (mean RMSE: 20.2; unpaired t-test, $p = 0.94$) and did not capture the discrimination responses of this neuron as well as the shape-dependent model (**Fig**. **3.5G**, mean RMSE: 18.4). This difference in performance was significant (unpaired t-test, $p < 0.001$) and specific to stimulus shape, as predicted response from the stimulus color-dependent model (**Fig. 3.5H**) was similar to the linear-gain model's prediction (mean RMSE: 20.2; unpaired t-test, $p = 0.73$). Thus, the shape-dependent gain model provided the best explanation of the change in responses displayed by neuron #2.

Observed responses of neuron #3 are shown in the bottom row panels of Figure 3.5 (solid lines and filled symbols). The linear-gain, polynomial-gain, and shape-dependent gain models (**Fig. 3.5I-K**) all produced similar discrimination response predictions for this neuron (mean RMSE: 12.4-12.5; unpaired t-test, $p > 0.3$ for all pairwise comparisons). However, the color-dependent model (**Fig. 3.5L**) explained the observed discrimination responses significantly better (mean RMSE = 11.9; unpaired t-test, $p<0.001$ for all pairwise comparisons to the other models). Thus, the color-dependent feature gain model provided the best explanation for this neuron's change in responses.

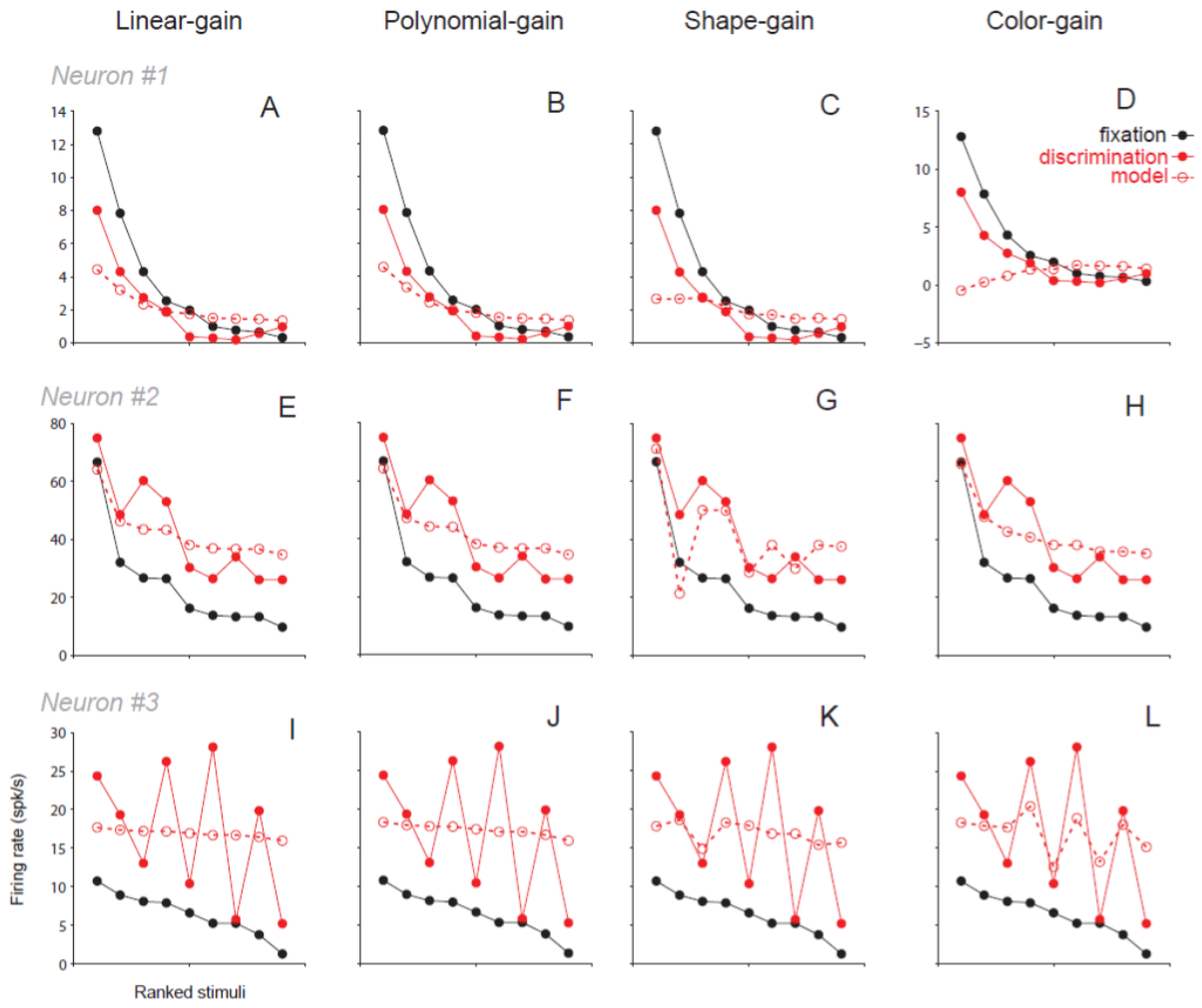


**Figure 3.5**. *Comparison of predicted model responses for example neurons #1-3.*

(A-D). Observed (red solid line and filled symbols, discrimination task responses) and predicted (red dashed line and open symbols) responses for regression models fit using fixation task responses of neuron #1 (black); discrimination and fixation responses are same data as Fig. 3.2B. The linear- (A) and polynomial gain models (B) have coefficients that depend on the firing rate of the neuron and produce predictions that match the observed data better than the stimulus-dependent predictions of the shape- and color-dependent gain models (C,D).

(E-H). Observed and predicted responses for regression models fit using fixation task responses of neuron #2; discrimination and fixation responses are same data as Fig. 3.3B. The model in (G) has separate scaling coefficients for stimuli with different shapes, and its prediction matches the observed data best. The rest of the models predict a more uniform scaling of the responses during the fixation task.

(I-L). Observed and predicted responses for regression models fit using fixation task responses of neuron #3; discrimination and fixation responses are same data as Fig. 3.4B. The model in (G) has separate scaling coefficients for stimuli with different colors, and its prediction matches the observed data best. The rest of the models predict a more uniform scaling of the responses during the fixation task.

### *3.3.3 Population response scaling*

We applied cross-validated RMSE to compare regression model performance for the entire population of 83 neurons. As above, we tested for significant differences between RMSE distributions for pairs of models ($\alpha = 0.01$) and compared mean RMSE values to determine the best-fitting regression model for each neuron.

For 18/83 neurons, the linear-gain model performed best out of the 4 models we considered. For 3/83 neurons, the polynomial-gain model performed best; thus, for 21/83 neurons in our population (25.3%), responses during discrimination could be expressed as a linearly or non-linearly scaled version of the responses during the fixation task.

For 20/83 neurons, the shape-dependent model performed best, and for 42/83 neurons the color-dependent model performed best. Thus, for 62/83 neurons in our population (74.7%), response scaling during discrimination was dictated by particular stimulus features.

### *3.3.4 Individual feature selectivity*

To understand whether the strength of selectivity for individual features related to the behavioral engagement, we examined shape and color selectivity for neurons that were best fit by the shape-, color-, or firing rate-dependent models (linear and polynomial). To obtain each neuron's selectivity for shape and color, we partitioned response variance due to stimulus shape and color (2-way ANOVA with interaction term; see Methods). The shape selectivity of the neuron was computed as the partitioned sum

of squares due to the effect of shape, normalized by the total sum of squares; color selectivity was computed equivalently.

We found a difference in the strength of feature selectivity among the neurons best fit by different models; selectivity during the discrimination task is shown in Figure 3.6 (results were similar for the fixation task). Shape-model neurons had significantly higher shape selectivity (**Fig**. **3.6A**, red) than color-model neurons (blue) and firing-rate model neurons (white; unpaired t-test, $p << 0.0001$). Color-model neurons and firing-rate model neurons had similar shape selectivity ($p = 0.14$).

The converse was true when comparing selectivity for color: neurons best explained by the color-dependent model (**Fig**. **3.6B**, blue) had significantly higher color selectivity than shape-model (red) and firing-rate model neurons (white; unpaired t-test, $p << 0.0001$). Shape-model and firing-rate model neurons had similar color selectivity (right panel, $p = 0.32$).

These results suggest that neurons whose difference in responses in the fixation and discrimination task is best captured with a dependence on a specific feature also have more of their response variance explained by that feature; and that the stimulus shape or color explains little of the variance in responses of neurons that are well-captured by the firing rate-dependent models.

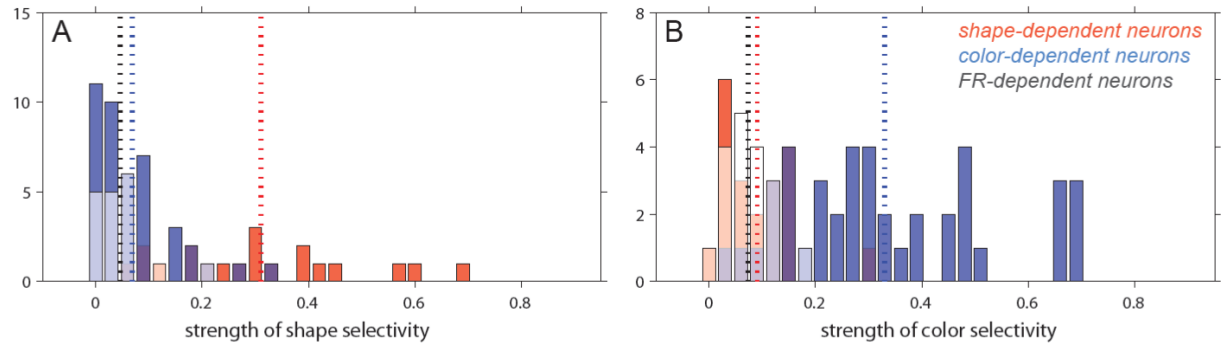


**Figure 3.6**. *Relationship between individual feature selectivity and best-fitting model.*

Individual feature selectivity was calculated as the partitioned variance due to shape or color from 2-way ANOVA (see Methods).

(A). Comparison between linear-gain and shape-dependent feature gain model. The histograms show the relative strength of shape selectivity for neurons whose responses are better fit by the shape-dependent feature gain model (red) compared to those that are better fit by the linear-gain model (white). Shape and color selectivity computed as partitioned variance from 2-way ANOVA (see Methods).

(B). Histograms as in (A), shown for relative strength of color selectivity in neurons whose responses are better fit by the color-dependent feature gain model (blue) compared to those that are better fit by the linear-gain model (white).

### *3.3.5 Relative feature selectivity*

Responses of V4 neurons carry information about both shape and color of objects, but in differing proportional amounts – that is, some neurons are more selective for shape and less for color, some the reverse, and some have similar selectivity for shape and color. To understand how behavioral engagement impacts this relative feature selectivity, we computed the ratio of shape selectivity to color selectivity, and compared this value for responses during the discrimination task (**Fig. 3.7**, ordinate) to the same ratio for responses during the fixation task (abscissa).

In this representation, the upper right quadrant contains neurons that are relatively more shape- than color-selective in both tasks. Neurons whose responses are best explained by the shape-dependent model (**Fig**. **3.7A**) are largely located in this quadrant (consistent with the individual feature selectivity findings in *3.3.4*). Many of these neurons fall displayed an increase in relative shape-selectivity during discrimination: many points fall above the identity line, and the change in selectivity in the two tasks was significant (paired t-test using log-transformed ratios, $p = 0.04$).

In contrast, the lower left quadrants in each panel of Figure 3.7 contain neurons that are relatively more color- than shape-selective in both tasks. Neurons whose responses are best explained by the color-dependent model (**Fig. 3.7B**) are located in this quadrant (consistent with the individual feature selectivity findings in *3.3.4*). Although many of these neurons became more color-selective during behavioral engagement (points falling below the identity line), this trend was not significant (paired t-test using log-transformed ratios, $p = 0.07$).

Neurons whose responses are best explained by the linear and non-linear firing-rate dependent models span across the two quadrants (**Fig**. **3.7C**), reflecting that responses of neurons in this group were not consistently more selective for either shape or color (see **Fig. 3.6**). Moreover, there was no significant change in relative feature selectivity in the two tasks (paired t-test using log-transformed ratios, $p = 0.15$).

These results show that responses of shape-dependent model neurons, specifically, exhibited an increase in relative shape selectivity during the shape discrimination task.

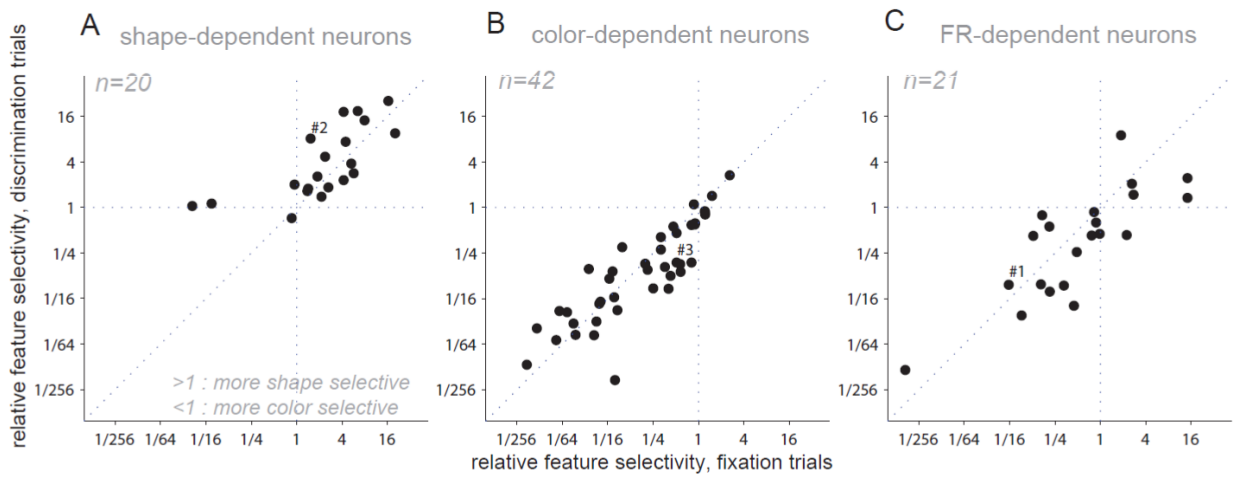


**Figure 3.7**. *Relationship between relative feature selectivity and best-fitting model.*

Relative feature selectivity, calculated as the ratio between partitioned variance due to shape and color from 2-way ANOVA (see Methods), is shown for discrimination trials (ordinate) compared to fixation trials (abscissa). Examples neurons #1-3 are marked.

(A). Relative feature selectivity for neurons whose responses were well-fit by the shape-dependent model. Relative selectivity of > 1 indicates most responses were more shape- than color-selective; neuronal responses were also relatively more shape-selective during discrimination than during fixation (paired t-test, $p = 0.04$).

(B). Relative feature selectivity for neurons whose responses were well-fit by the color-dependent model. Relative selectivity of < 1 indicates most responses were more color- than shape-selective; the selectivity of each neuron was similar during fixation and discrimination (paired t-test, $p = 0.07$).

(C). Relative feature selectivity for neurons whose responses were well-fit by the firing rate-dependent models. Responses of these neurons spanned the range from relatively more color-selective to relatively more shape-selective, and each neuron's selectivity was similar during fixation and discrimination (paired t-test, $p = 0.15$).

*3.3.6 Feature tuning in the population*

Next, we directly examined whether the feature tuning changed across the population of neurons in each group (best fit by the shape-, color-, or firing-rate-dependent model). To obtain tuning for stimulus shape, first, we computed shape-specific PSTHs by taking an average response for stimuli sharing a feature value (*e.g*. the 3 stimuli with the preferred shape). To obtain the population PSTH for shape, we then normalized and averaged shape-specific PSTHs for all neurons (see Methods for more details). Finally, we transformed the population shape-specific PSTH into a tuning curve by calculating the average response in the 50-200 ms window, as in Figs. 3.2B-3.4B. Since stimuli were tailored to each neuron and may have sampled different points along the individual tuning curves, we normalized the tuning curve relative to the response of the preferred shape. Color tuning curves were constructed equivalently, using color-specific responses.

We compared the average feature tuning curves for neurons whose responses are best explained by the shape-dependent model. The shape-selective tuning curve constructed responses during discrimination (red trace in **Fig. 3.8A**, left panel; bars show s.e.m.) was similar to the tuning curve constructed from responses during fixation (black trace). In contrast, the color tuning curve (**Fig. 3.8A**, right) was relatively flatter for discrimination responses than fixation responses, suggesting that tuning for color became comparatively broader with behavioral engagement. Average feature tuning in the neurons whose responses were governed by stimulus color (**Fig. 3.8B**) or firing rate during fixation (**Fig. 3.8C**) was similar for both features during the two tasks.

These results are consistent with the findings from the feature selectivity analyses (described in *3.3.4* and *3.3.5*), which suggested that shape-dependent neurons become relatively more shape-selective. Observations from the tuning curve analysis indicate that this significant change in feature selectivity comes about as a result of a loss of color selectivity, reflected in the broadening of color tuning in these neurons.

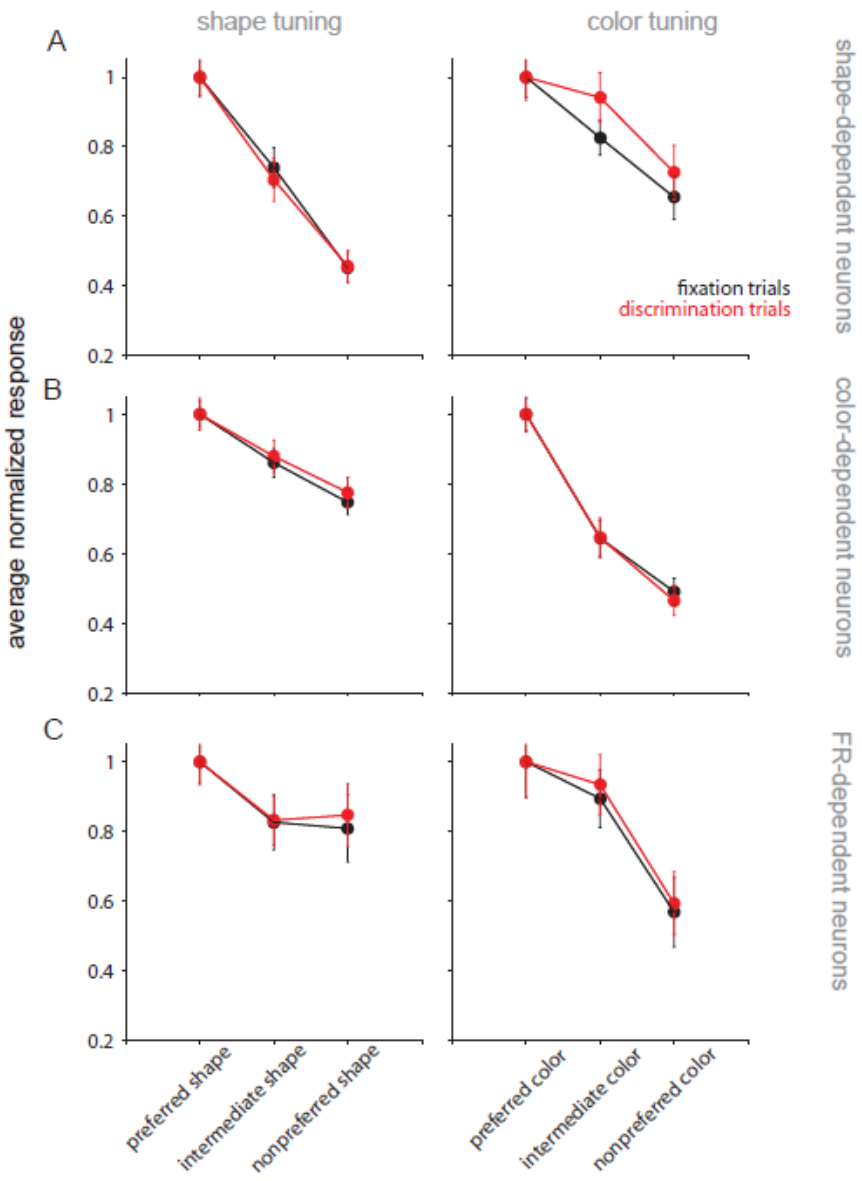


**Figure 3.8**. *Feature tuning in the population.*

Shape (left column) and color tuning curves (right column) were constructed by taking average responses (as in Fig. 3.2B) from shape- and color-specific PSTHs for each neuron. In this figure we include the shape and color tuning of neurons whose responses are best described by the shape-dependent model (A), color-dependent model (B), and firing-rate dependent models (C). Feature-specific PSTHs were computed by taking, *e.g.*, average responses for different shapes across all colors. Before extracting average responses for tuning curves, PSTHs were normalized for each task, and averaged across groups of neurons. Bars show s.e.m.

In 9 neurons from Animal 1, we were able to obtain not only data from the fixation and shape discrimination tasks, but also data from an equivalent task where animals judged stimulus color. Five neurons were those whose responses were best explained by the shape-dependent model, and 4 were best explained by the color-dependent model. The responses of all 9 neurons were similar between the two discrimination tasks, both in magnitude and selectivity. This result suggests that the modulations we observed in comparing responses during the fixation and shape discrimination task are not related to the behavioral relevance of the feature important for the discrimination task.

### *3.3.7 Influence of reference stimulus*

We considered whether other task-related variables could underlie the effects we observed: specifically, we wondered whether the relationship between the reference stimulus and neuronal preference, or the animal's saccadic eye movement, could account for the modulation in responses.

Previous studies of feature attention have reported that response modulation may be dictated by whether a neuron's feature selectivity/preference for object features matches the stimulus feature being attended (Treue and Martinez-Trujillo, 1999; also see Bichot *et al.*, 2005). For example, if a neuron's preferred shape is a square, and the reference stimulus in the discrimination task is also square, this neuron's responses would be enhanced; conversely, if a neuron's non-preferred shape is a circle, its responses would be suppressed when the reference stimulus is a square.

Although our study compares two task conditions that are different than those typically used to study attentional effects, we nonetheless examined whether this effect was present in our data. First, we compared the responses in discrimination trials where the reference stimulus matched (reference-match) or did not match (reference-mismatch) the preferred shape and color of each neuron, relative to the overall response in the discrimination trials (see Eqn. 3.4a). This relative difference was very small for most neurons; the median value was 0.09. For a neuron with a mean firing rate of 100 spk/s across the 9 stimuli, this value represents a mean stimulus response difference of <5 spk/s between the two datasets. The direction of this difference would be expected to be consistently positive, *i.e*. the reference-match trials should always be enhanced. The direction of response difference between reference-match and reference-mismatch trials was near 0 (median = -0.02; obtained by removing the absolute value operation from Eqn. 3.4a) and thus neurons in our population did not show a consistent enhancement of responses in reference-match trials.

Next, we examined whether tuning was similar in reference-match and reference-mismatch trials. We fit the linear and feature-dependent models separately for these subsets of data, and compared the regression coefficients excluding the constant regression terms (as these additive terms would not reflect changes in tuning). The linear-gain model produced very similar fits for reference-match and reference-mismatch trials in all neurons (paired t-test of regression coefficients from ref-match and ref-mismatch data subsets, $p = 0.6$; correlation between coefficients, $r = 0.99$, $p < 0.001$). The model fits from shape-dependent gain model also did not show a difference between reference-match and reference-mismatch trials (paired t-test, $p = 0.41$), although a more detailed

comparison suggested a slight change in tuning (correlation between ref-match and ref-mismatch for all coefficients, $r = 0.57$, $p < 0.001$; for preferred shape coefficient only: $r = 0.66$, $p < 0.001$). The model fits from the color-dependent model also showed no significant difference between ref-match and ref-mismatch trials (paired t-test, $p = 0.77$; correlation between ref-match and ref-mismatch for all coefficients, $r = 0.85$, $p < 0.001$). Thus, overall this result suggests that the relationship between reference stimulus identity and neuronal preference did not strongly contribute to the modulations we report above.

### *3.3.8 Influence of eye movement*

Finally, we assessed whether the animal's motor plan contributed to the modulation we observed. In correct trials where the reference stimulus and the test stimulus had the same shape, the animal always performed a saccade to the target location on the right side of the screen; when the two stimuli had different shapes, the saccadic movement was to the target location on the left. To determine whether this link between the stimuli and the eye movement demands of the task contributed to the effects we observed, we repeated the analyses above with our data separated into two groups based on the saccade outcome of the behavior task.

First, we examined whether the magnitude of the neuronal responses in rightward-saccade and leftward-saccade trials was relatively similar (see Eqn. 3.4b); the response difference was very small compared to overall response rate for most neurons (median = 0.05), equivalent to a <3 spk/s response difference for a neuron with a mean firing rate of 50 spk/s. Our recording sites in both animals accessed neurons with receptive field

locations in the lower portion of the right visual hemifield; this imbalance could be reflected in the difference between right- and left-ward saccade trials. We additionally examined the direction of the difference between trials ending in rightward and leftward saccades; this value was distributed around 0 in the population (median = -0.02; obtained by removing the absolute value operation from Eqn. 3.4b). Thus, although some individual neurons showed some difference in response related to saccade direction, across the whole population there was no difference in response magnitude.

Next, we assessed differences in tuning by performing the linear and feature-based gain regressions separately on these data subsets, and comparing the coefficients for rightward- and leftward-saccade trials. These values were highly consistent for both models, indicating no change in tuning: for the linear-gain model, the Pearson correlation coefficient between regression coefficients for rightward and leftward trials was $r = 0.999$ ($p < 0.001$); for the shape-dependent gain model, $r = 0.89$ (all coefficients, $p < 0.001$); for the color-dependent gain model, $r = 0.95$ (all coefficients, $p < 0.001$). These results suggest that eye movements and the corresponding motor plan did not significantly contribute to the modulatory effects we observed.

## 3.4 DISCUSSION

### *3.4.1 Summary of findings*

We compared responses of 83 V4 neurons to the 9 stimuli in two task contexts: a passive fixation and an active shape discrimination task. We found that most neurons displayed a change in response between the two tasks, and for 62/83 neurons in our population the modulation in responses during the discrimination task was best described on the basis of stimulus shape or color. Additionally, 3/83 neurons were well-fit by a nonlinear, FR-based gain model; only 18/83 neurons were best fit by a linear gain function. Responses of the 20 neurons that were best fit by the shape-dependent models also displayed a significant relative increase in shape selectivity during the discrimination task. This change, on average, appeared to be the result of broadened color tuning. Responses of 9/83 neurons for which the animal performed an additional color discrimination task were similar to responses of the same neurons in the shape discrimination task.

### *3.4.2 Relating observed trends to attention literature*

Our results showing changes in feature selectivity contrast with the broadly reported gain scaling of responses in sensory areas by attention (McAdams and Maunsell, 1999; Maunsell and Treue, 2006). The aim of our study was not to isolate attentional effects, but rather to compare visual encoding while an animal performs two tasks with different levels of behavioral engagement. Feature attention is likely to be at

play during the discrimination condition, since we ask the animals to discriminate stimulus shape while ignoring stimulus color, a behavioral goal similar to what has been used in other attentional and visual search studies (*e.g.* Mirabella *et al.*, 2007). Attentional mechanisms are also likely driving the responses of neurons that are firing-rate dependent (best fit by linear- and polynomial-gain models), as these neurons display response changes most consistent with reported results of attentional studies. The feature-dependent changes in response we observed in other neurons may be due to an additional mechanism which refines neuronal responses when the animal is engaged in a task, and which affects subpopulations of neurons that are more feature-selective. Such a mechanism may be established in the network linking V4 to higher-level cognitive areas such as the prefrontal cortex (Hamker, 2005; Gregoriou *et al.*, 2009; Liebe *et al.*, 2012), and is likely either developed or refined throughout the animals' training to perform the discrimination task.

### *3.4.3 Advantages of feature-dependent changes in response*

When responses are feature-dependent during behavioral engagement, relative feature selectivity changes. Responses of neurons best described by the shape-dependent model become relatively more shape-selective than color selective during behavior. Increased feature selectivity generally correlates with higher discriminability along the feature dimension; together, these results suggest that behavioral engagement may be increasing the discriminability of the dimension best captured by a neuron's responses (*e.g.* shape, for shape-selective neurons). This refinement of selectivity may

be accomplished by selectively strengthening feature inputs, such as by thresholding, to enhance the strongest, or preferred, inputs to a cell.

Moreover, in the subset of neurons whose responses during the discrimination task were shape-dependent, the increase in relative shape selectivity appeared to be correlated with broadened color tuning. Broadened tuning for a feature irrelevant for the discrimination is consistent with the enhanced selectivity for the relevant feature displayed by these neurons. Additionally, it is difficult to conclude whether tuning refinement observed in shape-dependent neurons improves shape coding without information about the noise covariance (Zhang and Sejnowski, 1999; Pouget *et al.*, 1999); further studies involving simultaneous recordings from neuronal populations may help resolve this uncertainty. In a few sessions where the animal performed both shape discrimination and color discrimination tasks, the modulation in neuronal responses relative to the fixation task was the same regardless of the type of discrimination. This observation suggests that the observed response changes may be due to the gross differences between fixation and discrimination tasks, and not the relevance of the object feature. However, the animal from which these additional data were obtained had been trained extensively to judge object shape, and only later trained to judge object color; thus, the task-related response modulation may in fact be affected by the animal's training history.

### *3.4.4 Relationship to other task-related factors*

An existing model of feature attention links non-uniform scaling effects to the relationship between neuronal selectivity and the attended (or "reference") stimulus in the

behavioral task (the “feature-similarity model”; Martinez-Trujillo and Treue, 1999, 2004; Bichot *et al*., 2005). This model predicts enhanced responses when the attended stimulus matched the preferred stimulus of the neuron, and suppressed responses when the attended stimulus did not match. This effect leads to an increased signal-to-noise ratio and non-multiplicative scaling of the population response, that is, a narrowing in the width of the population tuning curve. In our data, we found that our observed effects were not related to the feature similarity between the reference stimulus and the preference of the neuron.

In the sequential discrimination task, the animals could respond at will after the test stimulus was presented (a reaction-time task), at which time the stimulus was immediately extinguished; this was unlike the fixation task, where the animal maintained its gaze for a fixed duration of stimulus presentation. In addition, “same” and “different” shape judgments in the discrimination task were always linked to a rightward- and leftward-directed saccade, respectively. Together, the reaction-time and saccade direction elements of our task design may have introduced an effect of eye movement planning into the difference between the responses in the two tasks. Firstly, our analyses examined responses within a time window (50-200 ms from stimulus onset) well before the typical reaction time of the animals (250 ms). After our pilot recording sessions in this study, we also adjusted the presentation time of the stimuli during the fixation task to match the average presentation time of stimuli during the discrimination task (from 300 ms to 250 ms). Additionally, an influence of eye movement planning has been reported previously in V4: for example, when an animal’s eye movement targets a location inside the receptive field, responses increase, and outside the receptive field, responses

become suppressed (Han *et al*., 2009). Moreover, our recording sites were only able to target RFs in the lower portion of the visual hemifield, which could have differentially affected responses on trials when the animal made a saccade to the target on the same side as the RF versus the target in the opposite hemifield. We examined whether motor planning could explain our observed task-dependent effects by dividing our data into two subsets separated by the animal's saccade direction at the end of the trial. We found no relationship between eye movement and the response change during the discrimination task. This result may arise from two key differences: first, both of our targets lay outside the receptive field, so response changes consistent with previous findings would tend to be in the same direction and thus the magnitude of the difference would be smaller. Furthermore, we restricted our analysis window for neural responses to 50-200 ms from stimulus onset, and behavioral reaction times were ~260 ms. The response changes reported by Han *et al*. were typically within a 50 ms epoch before the saccadic eye movement, and thus would largely be absent in the portion of the response we chose to analyze.

# Chapter 4. Discussion and future directions

## *4.1 Summary and significance of findings*

In the preceding chapters, I have shown for the first time that the stimulus selectivity of V4 neurons depends on boundary orientation and object fill, and constructed a model that accounts for both attributes. These findings provide insights into how early visual signals may be combined to give rise to form selectivity in area V4: interior fill information is maintained in V4 neurons along with information about object boundaries, either in the form of edge phase or separate, unoriented signals. Responses of shape-selective V4 neurons reflect not only the shape of objects, but also their surface characteristics; this multiplexed representation suggests that responses in V4 could be used by downstream areas to parse complex natural scenes.

I have also shown that independent encoding of multiple object features is influenced by behavioral context; specifically, many shape- and color-selective neurons in area V4 are modulated in a feature-specific manner. During a shape discrimination task, the color tuning of the most shape-selective neurons broadened, suggesting that the effects of task context were related to the behavioral relevance of the feature dimension.

Together, these results provide both a bottom-up and a top-down perspective on responses in area V4 and illuminate the role that V4 neurons play in guiding visual behaviors such as object perception and discrimination.

*4.2 Directions for future investigations*

The results presented here demonstrated how single neurons in area V4 simultaneously encode multiple features of visual objects, and how this representation changes when the animal actively used information about the stimuli. To create a comprehensive understanding of how the brain processes visual objects and scenes, we must ultimately understand what information is contained in the complex, multidimensional representations in the ventral stream, and which computations extract the information relevant for perception.

To improve our understanding of how surface properties influence form representation in area V4, future experiments should examine how surface luminance gradients and texture modulate V4 responses to boundary form. Very few studies have examined selectivity for multiple features, and such experiments head towards a unifying framework for the disparate characterizations of V4 response properties. Additionally, it would be useful to examine how other network frameworks, such as deep learning nets, represent boundaries and interior characteristics of objects, and whether they produce similar responses to filled and outline stimuli. Although responses of neurons in areas V1-V4 tend to reflect object parts, ultimately our perception is holistic. Learning how neurons build up multiplexed representations will go a long way towards understanding how we see objects as a whole, yet are also able to judge and compare their components individually.

Rapidly changing behavioral goals impose substantial demands on the primate visual brain. Its organization and information processing flow reflect a system constrained by time, the need for accuracy, and resource availability. To further decrypt how the brain

performs visual tasks, future experiments should examine neural responses in more challenging, life-like situations. For example, it would be important to understand how the brain encodes visual information in conditions where objects are partially obscured or otherwise less salient, such as in crowded scenes, or when there are competing task demands, such as tasks where a dynamic reward scheme changes which object parts are most relevant. It would also be important to perform simultaneous multi-area recording, *e.g.* in IT or prefrontal cortex, and analysis of response dynamics during behavior. Such an experiment could provide a greater insight into how information may be decoded by downstream areas, and may identify the source of behavior-related response modulation.

Lastly, the computational models we developed were able to account for our experimental observations, but more studies would be necessary to produce a more general model: for example, the current model does not reflect information about color and luminance that is present in V4 responses. Additionally, this model is purely feedforward (lacking retrograde connections) and thus cannot reflect top-down influences, which are important for feature encoding during behavior. The model also lacks representation of time: input and output represent mean responses to stimuli, lacking the rich information contained in the dynamics of V4 responses. A collaboration between experimental and computational efforts will produce a more elaborate picture of how V4 feature selectivity emerges, how it changes in time, and what other brain areas provide the strongest influence to produce the dynamic and varied responses in area V4.

## REFERENCES


Albrecht, D. G., Valois, R. L. De, & Thorell, L. G. (1980). Visual Cortical Neurons: Are Bars or Gratings the Optimal Stimuli? *Science*, *207*(4426), 88–90. https://doi.org/10.2307/1683218

Allen, M., Bloom, P., & Hodgson, E. (2010). Do Young Children Know What Makes A Picture Useful To Other People? *Journal of Cognition and Culture*, *10*(1), 27–37. https://doi.org/10.1163/156853710X497158

Bichot, N. P., Heard, M. T., DeGennaro, E. M., & Desimone, R. (2015). A Source for Feature-Based Attention in the Prefrontal Cortex. *Neuron*, *88*(4), 832–844. https://doi.org/10.1016/j.neuron.2015.10.001

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, *20*(1), 38–64. https://doi.org/10.1016/0010-0285(88)90024-2

Brodmann, K. (1909). *The principles of comparative localisation in the cerebral cortex based on cytoarchitectonics*. Leipzig: Barth. Retrieved from http://digital.zbmed.de/zbmed/id/554966

Buffalo, E. A., Fries, P., Landman, R., Liang, H., & Desimone, R. (2010). A backward progression of attentional effects in the ventral stream. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(1), 361–5. https://doi.org/10.1073/pnas.0907658106

Bushnell, B. N., Harding, P. J., Kosai, Y., Bair, W., & Pasupathy, A. (2011). Equiluminance cells in visual cortical area v4. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(35), 12398–412. https://doi.org/10.1523/JNEUROSCI.1890-11.2011

Bushnell, B. N., Harding, P. J., Kosai, Y., & Pasupathy, A. (2011). Partial occlusion modulates contour-based shape encoding in primate area V4. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(11), 4012–24. https://doi.org/10.1523/JNEUROSCI.4766-10.2011

Bushnell, B. N., & Pasupathy, A. (2012). Shape encoding consistency across colors in primate V4. *Journal of Neurophysiology*, *108*(5), 1299–308. https://doi.org/10.1152/jn.01063.2011

Cadieu, C., Kouh, M., Pasupathy, A., Connor, C. E., Riesenhuber, M., & Poggio, T. (2007). A model of V4 shape selectivity and invariance. *Journal of Neurophysiology*, *98*(3), 1733–50. https://doi.org/10.1152/jn.01265.2006

Cloutman, L. L. (2013). Interaction between dorsal and ventral processing streams: Where, when and how? *Brain and Language*, *127*(2), 251–263. https://doi.org/https://doi.org/10.1016/j.bandl.2012.08.003

Conway, B. R., Moeller, S., & Tsao, D. Y. (2007). Specialized Color Modules in Macaque Extrastriate Cortex. *Neuron*, *56*(3), 560–573. https://doi.org/10.1016/j.neuron.2007.10.008

David, S. V., Hayden, B. Y., & Gallant, J. L. (2006). Spectral Receptive Field Properties Explain Shape Selectivity in Area V4. *Journal of Neurophysiology*, *96*(6). Retrieved from http://jn.physiology.org/content/96/6/3492.short

David, S. V, Hayden, B. Y., Mazer, J. A., & Gallant, J. L. (2008). Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron*, *59*(3), 509–521.

De Weerd, P., Desimone, R., & Ungerleider, L. G. (1996). Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Visual Neuroscience*, *13*(3), 529. https://doi.org/10.1017/S0952523800008208

Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, *57*(3). Retrieved from http://jn.physiology.org/content/57/3/835.short

El-Shamayleh, Y., & Pasupathy, A. (2016). Contour Curvature As an Invariant Code for Objects in Visual Area V4. *Journal of Neuroscience*, *36*(20). Retrieved from http://www.jneurosci.org/content/36/20/5532.full

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1822724

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), 974–81. https://doi.org/10.1038/nn.3402

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W., & Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, *76*(4). Retrieved from http://jn.physiology.org/content/76/4/2718.short

Gallant, J. L., Shoup, R. E., & Mazer, J. A. (2000). A Human Extrastriate Area Functionally Homologous to Macaque V4. *Neuron*, *27*(2), 227–235. https://doi.org/10.1016/S0896-6273(00)00032-5

Gattass, R., Sousa, a P., & Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *8*(6), 1831–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3385477

Goda, N., Tachibana, A., Okazawa, G., & Komatsu, H. (2014). Representation of the Material Properties of Objects in the Visual Cortex of Nonhuman Primates. *Journal of Neuroscience*, *34*(7). Retrieved from http://www.jneurosci.org/content/34/7/2660

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci*, *15*(1), 20–25.

Goodale, M. A., & Westwood, D. A. (2004). An evolving view of duplex vision: separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology*, *14*(2), 203–211. https://doi.org/https://doi.org/10.1016/j.conb.2004.03.002

Gregoriou, G. G., Gotts, S. J., Zhou, H., & Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, *324*(5931), 1207–10. https://doi.org/10.1126/science.1171402

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*(2), 173–211. https://doi.org/10.1037/0033-295X.92.2.173

Haenny, P. E., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. I. Single cell activity in V1 and V4 on visual tasks. *Experimental Brain Research*, *69*(2), 225–44. https://doi.org/10.1007/BF00247569

Hamker, F. H. (2005). The Reentry Hypothesis: The Putative Interaction of the Frontal Eye Field, Ventrolateral Prefrontal Cortex, and Areas V4, IT for Attention and Eye Movement. *Cerebral Cortex*, *15*(4), 431–447. https://doi.org/10.1093/cercor/bhh146

Hansen, T., & Gegenfurtner, K. R. (2007). Higher Order Color Mechanisms for Image Segmentation. In F. Mele, G. Ramella, S. Santillo, & F. Ventriglia (Eds.), *Advances in Brain, Vision, and Artificial Intelligence: Second International Symposium, BVAI 2007, Naples, Italy, October 10-12, 2007. Proceedings* (pp. 72–83). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-75555-5_8

Heywood, C., & Cowey, A. (1987). On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. *Journal of Neuroscience*, *7*(9), 2601–2617. Retrieved from http://www.jneurosci.org/content/7/9/2601.short

Hochberg, J., & Brooks, V. (1962). Pictorial Recognition as an Unlearned Ability: A Study of One Child's Performance. *The American Journal of Psychology*, *75*(4), 624. https://doi.org/10.2307/1420286

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, *148*(3), 574–591. https://doi.org/10.1113/jphysiol.1959.sp006308

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

Ibos, G., & Freedman, D. J. (2014). Dynamic Integration of Task-Relevant Visual Features in Posterior Parietal Cortex. *Neuron*, *83*(6), 1468–80. https://doi.org/10.1016/j.neuron.2014.08.020

Ito, M., Fujita, I., Tamura, H., & Tanaka, K. (1994). Processing of Contrast Polarity of Visual Images in Inferotemporal Cortex of the Macaque Monkey. *Cerebral Cortex*, *4*(5), 499–508. https://doi.org/10.1093/cercor/4.5.499

Ito, M., & Komatsu, H. (2004). Representation of Angles Embedded within Contour Stimuli in Area V2 of Macaque Monkeys. *The Journal of Neuroscience*, *24*(13), 3313–3324. https://doi.org/10.1523/jneurosci.4364-03.2004

Kayaert, G., Biederman, I., & Vogels, R. (2003). Shape Tuning in Macaque Inferior Temporal Cortex. *Journal of Neuroscience*, *23*(7). Retrieved from http://www.jneurosci.org/content/23/7/3016.full

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*(3). Retrieved from http://jn.physiology.org/content/71/3/856.short

Kosai, Y., El-Shamayleh, Y., Fyall, A. M., & Pasupathy, A. (2014). The Role of Visual Area V4 in the Discrimination of Partially Occluded Shapes. *Journal of Neuroscience*, *34*(25). Retrieved from http://www.jneurosci.org/content/34/25/8570.short

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*(1), 26–49. https://doi.org/10.1016/j.tics.2012.10.011

Latham, K., & Whitaker, D. (1996). A Comparison of Word Recognition and Reading Performance in Foveal and Peripheral Vision. *Vision Research*, *36*(17), 2665–2674. https://doi.org/10.1016/0042-6989(96)00022-3

Leung, T., & Malik, J. (1998). Contour continuity in region based image segmentation. In H. Burkhardt & B. Neumann (Eds.), *Computer Vision --- ECCV'98: 5th European Conference on Computer Vision Freiburg, Germany, June, 2--6, 1998 Proceedings, Volume I* (pp. 544–559). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/BFb0055689

Levitt, J. B., Kiper, D. C., & Movshon, J. A. (1994). Receptive fields and functional architecture of macaque V2. *Journal of Neurophysiology*, *71*(6), 2517–2542. Retrieved from http://jn.physiology.org/content/71/6/2517.short

Liebe, S., Hoerzer, G. M., Logothetis, N. K., & Rainer, G. (2012). Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nature Neuroscience*, *15*(3), 456–462. https://doi.org/10.1038/nn.3038

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, *77*(1), 24–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9120566

Martinez-Trujillo, J., & 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

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–22. https://doi.org/10.1016/j.tins.2006.04.001

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. *Journal of Neuroscience*, *19*(1). Retrieved from http://www.jneurosci.org/content/19/1/431

McAdams, C., & Maunsell, J. (2000). Attention to both space and feature modulates neuronal responses in macaque area V4. *Journal of Neurophysiology*, 1751–1755. Retrieved from http://jn.physiology.org/content/83/3/1751.short

Mechler, F., & Ringach, D. L. (2002). On the classification of simple and complex cells. *Vision Research*, *42*(8), 1017–1033. https://doi.org/10.1016/S0042-6989(02)00025-1

Mel, B. W. (1997). SEEMORE: Combining Color, Shape, and Texture Histogramming in a Neurally Inspired Approach to Visual Object Recognition. *Neural Computation*, *9*(4), 777–804. https://doi.org/10.1162/neco.1997.9.4.777

Merigan, W. H. (1996). Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. *Visual Neuroscience*, *13*(1), 51–60. https://doi.org/10.1017/S0952523800007124

Merigan, W. H., & Pham, H. A. (1998). V4 lesions in macaques affect both single- and multiple-viewpoint shape discriminations. *Visual Neuroscience*, *15*, 359–367. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/05E103978D4D3EC67D50A0DD7B661BEA/S0952523898152112a.pdf/v4_lesions_in_macaques_affect_both_single_and_multipleviewpoint_shape_discriminations.pdf

Mirabella, G., Bertini, G., Samengo, I., Kilavik, B. E., Frilli, D., Della Libera, C., & Chelazzi, L. (2007). Neurons in area V4 of the macaque translate attended visual features into behaviorally relevant categories. *Neuron*, *54*(2), 303–18. https://doi.org/10.1016/j.neuron.2007.04.007

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782–784. Retrieved from https://books.google.com/books?hl=en&lr=&id=YwSKMQIymxsC&oi=fnd&pg=PA342&dq=+v4+attention&ots=V7E5G0OOKH&sig=aZ4OUA1K0dKYg8y2DRu0O2xFWSk#v=onepage&q=v4 attention&f=false

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), 909–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8229178

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978a). Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *The Journal of Physiology*, *283*(1), 101–120. https://doi.org/10.1113/jphysiol.1978.sp012490

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978b). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *The Journal of Physiology*, *283*(1), 53–77. https://doi.org/10.1113/jphysiol.1978.sp012488

Mumford, D., Kosslyn, S. M., Hillger, L. A., & Herrnstein, R. J. (1987). Discriminating figure from ground: The role of edge detection and region growing. *Proceedings of the National Academy of Sciences of the United States of America*, *84*, 7354–7358. Retrieved from http://www.pnas.org/content/84/20/7354.long

Namima, T., Yasuda, M., Banno, T., Okazawa, G., & Komatsu, H. (2014). Effects of luminance contrast on the color selectivity of neurons in the macaque area v4 and inferior temporal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(45), 14934–47. https://doi.org/10.1523/JNEUROSCI.2289-14.2014

Ogawa, T., & Komatsu, H. (2004). Target Selection in Area V4 during a Multidimensional Visual Search Task. *The Journal of Neuroscience*, *24*(28), 6371–6382. https://doi.org/10.1523/jneurosci.0569-04.2004

Okazawa, G., Tajima, S., & Komatsu, H. (2015). Image statistics underlying natural texture selectivity of neurons in macaque V4. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(4), E351-60. https://doi.org/10.1073/pnas.1415146112

Oleskiw, T. D., Pasupathy, A., & Bair, W. (2014). Spectral receptive fields do not explain tuning for boundary curvature in V4. *Journal of Neurophysiology*, *112*(9). Retrieved from http://jn.physiology.org/content/112/9/2114

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews. Neuroscience*, *8*(5), 379–91. https://doi.org/10.1038/nrn2131

Pasupathy, A., & Connor, C. E. (1999). Responses to Contour Features in Macaque Area V4. *J Neurophysiol*, *82*(5), 2490–2502. Retrieved from http://jn.physiology.org/content/82/5/2490.full

Pasupathy, A., & Connor, C. E. (2001). Shape Representation in Area V4: Position-Specific Tuning for Boundary Conformation. *J Neurophysiol*, *86*(5), 2505–2519. Retrieved from http://jn.physiology.org/content/86/5/2505.full

Perkel, D. J., Bullier, J., & Kennedy, H. (1986). Topography of the afferent connectivity of area 17 in the macaque monkey: A double-labelling study. *The Journal of Comparative Neurology*, *253*(3), 374–402. https://doi.org/10.1002/cne.902530307

Pinna, B., & Grossberg, S. (2005). The watercolor illusion and neon color spreading: a unified analysis of new cases and neural mechanisms. *Journal of the Optical Society of America A*, *22*(10), 2207–2221. https://doi.org/10.1364/JOSAA.22.002207

Pouget, A., Deneve, S., Ducom, J.-C., & Latham, P. E. (1999). Narrow Versus Wide Tuning Curves: What's Best for a Population Code? *Neural Computation*, *11*, 85–90. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.99.4723&rep=rep1&type=pdf

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive Mechanisms Subserve Attention in Macaque Areas V2 and V4. *J. Neurosci.*, *19*(5), 1736–1753. Retrieved from http://www.jneurosci.org/content/19/5/1736.full

Reynolds, J., 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

Rodríguez-Sánchez, A. J., & Tsotsos, J. K. (2012). The Roles of Endstopped and Curvature Tuned Computations in a Hierarchical Representation of 2D Shape. *PLoS ONE*, *7*(8), e42058. https://doi.org/10.1371/journal.pone.0042058

Rosenholtz, R. (2016). Capabilities and Limitations of Peripheral Vision. *Annual Review of Vision Science*, *2*(1), 437–457. https://doi.org/10.1146/annurev-vision-082114-035733

Rust, N., & DiCarlo, J. (2010). Selectivity and Tolerance ("Invariance") Both Increase as Visual Information Propagates from Cortical Area V4 to IT. *The Journal of Neuroscience*, *30*(39), 12978–12995. https://doi.org/10.1523/jneurosci.0179-10.2010

Schein, S. J., & Desimone, R. (1990). Spectral properties of V4 neurons in the macaque. *J Neurosci*, *10*(10), 3369–3389.

Schiller, P. H. (1995). Effect of lesions in visual cortical area V4 on the recognition of transformed objects. *Nature*, *376*(6538), 342–344. https://doi.org/10.1038/376342a0

Schiller, P., & Lee, K. (1991). The role of the primate extrastriate area V4 in vision. *Science*, *251*(4998), 1251–1253. https://doi.org/10.1126/science.2006413

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology*, *79*(6), 3272–3278.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, *240*(4850), 338–340. https://doi.org/10.1126/science.3353728

Tootell, R., Silverman, M., Hamilton, S., Switkes, E., & De Valois, R. (1988). Functional anatomy of macaque striate cortex. V. Spatial frequency. *Journal of Neuroscience*, *8*(5). Retrieved from http://www.jneurosci.org/content/8/5/1610.short

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

Van Essen, D. C., & Gallant, J. L. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron*, *13*(1), 1–10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8043270

von der Heydt, R., & Pierson, R. (2006). Dissociation of color and figure–ground effects in the watercolor illusion. *Spatial Vision*, *19*(2), 323–340. https://doi.org/10.1163/156856806776923416

Wagemans, J., De Winter, J., de Beeck, H. O., Ploeger, A., Beckers, T., & Vanroose, P. (2008). Identification of Everyday Objects on the Basis of Silhouette and Outline Versions. *Perception*, *37*(2), 207–244. https://doi.org/10.1068/p5825

Zamarashkina, P., Popovkina, D., & Pasupathy, A. (2017). Stimulus and task dependence of response latencies in primate area V4. *Journal of Vision*, *17*(10), 476.

Zeki, S. (1980). The representation of colours in the cerebral cortex. *Nature*, *284*(5755), 412–418. https://doi.org/10.1038/284412a0

Zeki, S. M. (1973). Colour coding in rhesus monkey prestriate cortex. *Brain Research*, *53*, 422–427. Retrieved from http://www.vislab.ucl.ac.uk/pdf/colour_coding_in_monkey_prestriate.pdf

Zhang, K., & Sejnowski, T. J. (1999). Neuronal Tuning: To Sharpen or Broaden? *Neural Computation*, *11*(1), 75–84. https://doi.org/10.1162/089976699300016809

Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *The Journal of Neuroscience*, *20*(17), 6594–611. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10964965

Ziemba, C. M., Freeman, J., Movshon, J. A., & Simoncelli, E. P. (2016). Selectivity and tolerance for visual texture in macaque V2. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(22), E3140-9. https://doi.org/10.1073/pnas.1510847113