Neural Representation of Surface Ordering in Visual Areas V1, V2 and MT

Thesis by
Grace C. Chang

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, CA
2003
(Defended May 20, 2003)

Acknowledgements

At the Salk Institute:

Thomas Albright, Gene Stoner, Jennifer Costanza, Dinh-Dinh Diep, Laura Wilson, Kerri Sevenbergen, Margaret Mitchell, Trina Norden-Krichmar, the computer support staff.

At Caltech:

John Allman, Christof Koch, Mark Konishi, Pietro Perona, Michael Hoffmann.

At Stanford:

Denis Baylor.

Above all, I am grateful for the friendships which have nurtured me during my extended graduate career. This thesis would not have come to fruition without the kind, thoughtful, and generous emotional support of many individuals. I would particularly like to thank the following: Jennifer Costanza, Sunil Gandhi, Dinh-Dinh Diep, Amit Manwani, Laurent Itti, Giacomo Indiveri, Timmer Horiuchi, Jennifer Linden, Daniel Maskit, Suzanne Biegel, Frank Perez, Dan Ruderman, Kellie Wills, Sean Nhan, Sharon Horton, Robert Liu, Robert Drost, Sharon Drost, Sherleen Huang, Sarah Hung, Lisa Marin, Peggy Wu, Jake Brenner, Pamela Mosher, my other medical school classmates, and the CTY/Mattababy crowd. While a list of this nature does not do justice to the unique wonders these individuals possess, I hope they will forgive me, knowing how much I appreciate the extraordinary qualities in each. I am indebted to them and to the many others who have offered me their warm hearts in friendship.

I'd like to additionally mention some august faculty members who influenced me greatly. At Caltech, Professor Joel Franklin was the source of much comfort and learned advice, which I continue to carry with me. Dean Michael Hoffmann offered significant hope and perspective. Finally, Professor (now Dean) James Plummer and Professor Denis Baylor at Stanford unwittingly set the standard for my scientific endeavors during my pre-Caltech years, through their unparalleled talent, integrity, modesty and insight. I am grateful and privileged to have received much additional wisdom and support from Denis Baylor since.

For all these relationships which have inspired me and sustained my spirit, I am truly appreciative.

ABSTRACT

Visual cortical areas V1, V2 and MT may participate in the representation of surface ordering, the arrangement of one surface in front of another. This work investigates the role of neurons of V1 and V2 in figure-ground representation in static stimuli, as well as the role of MT in surface ordering in dynamic stimuli.

Electrical recordings were made in V1 and V2 to determine whether neurons in these areas encode information about the identities of figure and ground, and also whether they respond to figure-ground cues. We recorded from 3 monkeys, one trained on a fixation task, and the other two on a match-to-sample task that ensured attention to the stimuli. The stimuli consisted of rectangles of differing contrast arranged in an unambiguous or ambiguous figure-ground configuration. The stimuli were positioned so that the cells' receptive fields were located either at the border between rectangles or in the interiors of rectangles. Cells demonstrating selectivity at borders or interiors of unambiguous figure-ground stimuli were considered selective for border ownership or figure vs. ground, respectively. Cells showing selectivity at borders or interiors of ambiguous figure-ground stimuli were considered selective for figure-ground cues.

Preliminary experiments on the fixating monkey suggested that a small fraction of cells in V1 and V2 might play a role in figure-ground interpretation. The results from the awake, behaving monkeys further support the hypothesis that V1 and V2 play a role in figure-ground perception. In both areas we found cells demonstrating selectivity for border ownership, and in V2 we found cells demonstrating selectivity for figure over ground. However, in V1 and V2 there was also evidence a separate population was responding to the presence of figure-ground cues in the stimulus.

The experiments in MT were performed on two awake, behaving monkeys. The stimuli were transparent rotating cylinders comprised of random dots moving along a sinusoidal gradient. The stimuli were bistable—perceived to rotate in one direction or its opposite. The monkeys indicated in which direction they perceived the cylinder's front surface rotating. Cells were found whose firing correlated with the monkeys' bistable percept, even though the stimuli were identical.

CONTENTS

LIST OF FIGURES

PART I

Figure-Ground Representation in V1 and V2

CHAPTER 1

Abstract:

Previous studies by others have suggested that neurons in visual cortical areas V1 and V2 may participate in figure-ground interpretation. The purpose of this work is to investigate whether responses of neurons in these areas indeed encode information about the identities of figure and ground or whether the responses are modulated by figure-ground cues. We observed the responses of neurons in V1 and V2 to stimuli that consisted of a pair of overlapping rectangles with different contrast. The stimuli were designed to reveal two types of selectivity that may be involved in figure-ground computations: a) "border ownership" – differential responses to an inter-rectangle border depending on which side of the border the foreground rectangle lies, and b) "figure-ground selectivity" – differential responses to a figure or ground rectangle when the cell's classical receptive field lies completely within the rectangle. The T-junctions that determine figure-ground context in the stimuli were always positioned outside the neurons' classical receptive fields. To control for responsiveness to incidental features of the stimuli, or to the presence of T-junctions per se, we used stimuli with ambiguous figure-ground cues.

Preliminary experiments on one monkey trained to fixate on the stimuli suggested that a small fraction of cells in V1 and V2 may play a role in figure-ground interpretation. In V1, one cell of 20 tested exhibited weak border ownership selectivity, and three showed weak figure-ground selectivity at the better of two positions for the figure. In V2, five of 51 cells showed border ownership selectivity and eight showed figure-ground

selectivity at the better of two figure positions. With ambiguous figure-ground stimuli, only one of 20 cells in V1 responded preferentially to the borders between the rectangles, and no cell responded selectively to the interior regions of the rectangles. In V2, one of 51 cells responded preferentially to the borders, and 11 of the 51 responded preferentially to the interiors of these ambiguous figure-ground stimuli.

To pursue these observations, we studied neuronal responses in two additional monkeys trained to perform a match-to-sample task, requiring attention to figure-ground cues in the stimuli. In V1, 10 of 103 cells showed weak border ownership selectivity in the unambiguous figure-ground stimuli, while in V2 36 of 170 cells showed similar behavior. A different group of cells responded selectively to the borders of ambiguous figure-ground stimuli, suggesting that a separate group of cells responds selectively to figure-ground cues rather than to border ownership. These cells comprised 6 of 97 cells studied in V1 and 16 of 152 cells in V2. In V1, there was little evidence that cells responded selectively to the interior region of unambiguous figure-ground or ambiguous figure-ground stimuli. Thus, given the criteria for selectivity, sampling of the roughly 100 cells studied would have been expected to yield apparently selective cells with about the frequency observed. In V2, there was more evidence for a population of figure-ground selective cells. 17 of 170 cells demonstrated selectivity in the interior region of true figure-ground stimuli, and a large majority of these preferred figure over ground. 19 of 152 other cells were selective for the interiors of ambiguous figure-ground stimuli. This points to the existence of a population that responds to figure-ground cues.

These results further support a role for V1 and V2 play a role in figure-ground perception. In both areas we found cells which demonstrated selectivity for border

ownership, and in V2 we found cells which demonstrated selectivity for figure over ground. However, in both V1 and V2 there was evidence that a separate population responded to the presence of figure-ground cues in the stimulus.

Introduction:

Our ability to evaluate the structure of the world around us and to navigate through it depends critically on our visual system's ability to separate figure (what lies in the foreground) from ground (what lies in the background). Despite the importance of this computation, the neural mechanisms that mediate it remain largely unknown. Early studies in primate visual cortex looked only at isolated retinal image features, such as motion, color, and orientation, within the classical receptive fields of neurons (Hubel and Wiesel 1968, 1977). However, these studies could not address the problem of how we are able to segregate figure from ground. For example, in Fig 1A, when the classical receptive field is located at the border between two sets of rectangles, the local image features within the classical receptive fields are identical, but our interpretation of the figural aspects of the stimuli are very different. In the left-hand panel, the figure lies to the left, whereas in the right-hand panel, the figure lies to the right. Therefore, studies of purely local image properties could not reveal "border ownership," i.e., to which side of a border the figure lies. Similarly, in Fig 1B, the classical receptive field is located in interior regions of the rectangles, the local image features are identical—but in one instance the region is the interior of a figure rectangle, whereas in the other, the region is the interior of a ground rectangle. Again, classical receptive field properties cannot explain figure-ground discriminations within the interior regions of visual stimuli.

FIGURE 1

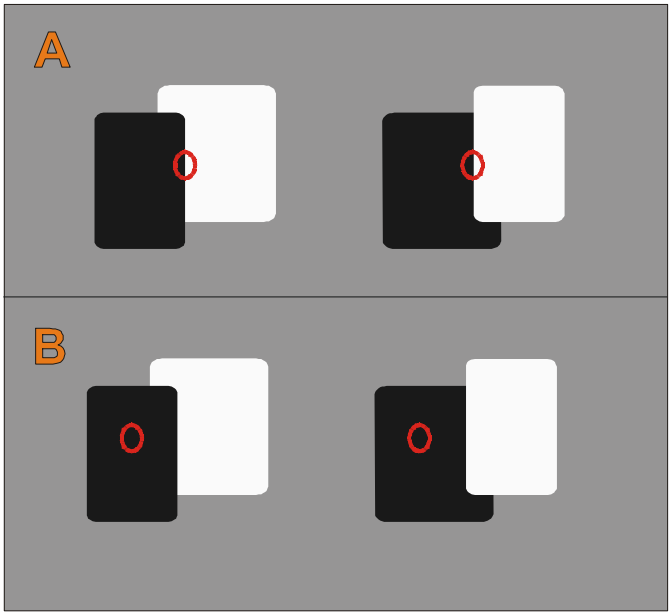

Figure 1: The figure-ground problem. The red ovals in these diagrams indicate classical receptive field boundaries. The figure rectangle is considered the rectangle in front, the ground rectangle is considered the rectangle behind. A: The classical receptive field is at the border between rectangles in two separate cases. The pair of rectangles on the left has a dark rectangle in front of a light rectangle, whereas the pairon the right has a dark rectangle behind a light rectangle. The local image features within the classical receptive fields are identical in both sets of rectangles, but the figure-ground direction in each case is different. For the rectangles on the left, the figure lies to the left, whereas for the rectangles on the right, the figure lies to the right. B: The classical receptive field is contained within the interior of pairs of rectangles. Again, the local image properties are identical within the classical receptive fields. However, in the pair of rectangles on the left, the classical receptive field is contained within the figure rectangle, whereas in the pair of rectangles on the right, the classical receptive field is contained within the ground rectangle.

Therefore, while local image properties contained within a classical receptive field are important to parsing some basic aspects of the visual scene, it is the surrounding context which guides our visual system to a solution of the figure-ground problem.

Indeed, several studies have shown that the responses of many neurons to stimuli within the classical receptive field are influenced by stimuli located in the larger, surrounding areas (Kaffei and Fiorentini, 1976; Nelson and Frost, 1978; Allman et al., 1985; Gilbert and Wiesel, 1990; Knierim and Van Essen, 1992; Sillito and Jones, 1996; Li et al., 2000; Li et al., 2001). It has been suggested that this contextual modulation is involved in such properties of global visual processing as feature contrast, which may be necessary for feature discrimination and visual search (Allman et al., 1985; Knierim and Van Essen, 1992; Kastner et al., 1997), illusory contour detection (von der Heydt et al., 1984; Peterhans and von der Heydt, 1989), and surface perception (Rossi et al., 1996; MacEvoy et al., 1998; Huang et al., 2002). Additional studies have suggested that contextual modulation may also play a role in figure-ground interpretation (von der Heydt et al., 1993; Lamme 1995; Zipser et al., 1996; Baumann et al., 1997; Heitger et al., 1998; Lee et al., 1998; Zhou et al., 2000). Specifically, both Baumann et al. (1997) and Zhou et al. (2000) reported cells with border ownership selectivity at contour borders between figure and ground regions in area V2 in awake, fixating monkeys. However, unlike Baumann et al., Zhou et al. also reported similar findings in V1. Furthermore, studies of area V1 in awake fixating monkeys by Lamme (1995), Zipser at al. (1996), and Lee et al. (1998) found figure vs. ground selectivity when receptive fields were located entirely within the interiors of figure regions.

The visual stimuli used in all these experiments varied: Baumann et al. (1997) used occluding contour stimuli, in which light or dark rectangles overlay line-gratings with the opposite contrast. They found neurons in V2 which responded to border ownership, indicating to which side of the occlusion border the figure lay. They did not find neurons in V1 which behaved in this way. In a more recent study, Zhou et al. (2000) tested single light or dark solid squares on solid backgrounds, as well as light and dark overlapping squares on solid backgrounds, to examine cells' preferences for border ownership. They also reported neurons in V2, but additionally in V1, which seemed to encode border ownership at the contrast-defined edges of these squares.

Other studies have compared responses of V1 neurons, in particular, to figure vs. ground when their classical receptive fields were positioned within the interiors of figure and ground regions. Zipser et al. (1996) reported enhanced V1 responses to a textured, oriented figure on a similarly textured background with different orientation. However, a follow-up study by Rossi et al. (2001) suggested that V1 neurons appear to give selectively enhanced responses to texture boundaries located near the edge of the classical receptive field, rather than to the interiors of figure regions per se.

Suggestive as they are, these previous studies suffer from several limitations. First, it is not clear whether the monkeys were truly attending to the spatial configurations of the stimuli; these studies were performed on either fixating monkeys or on monkeys that were trained to detect a stimulus rather than to discriminate between two stimuli. Second, it is not clear whether neurons in V1 and V2 differ in border ownership selectivity or true figure-ground selectivity. Third, it is not clear whether the cells were responding to true figure-ground context or rather to figure-ground cues, such as T-

junctions, which could be used for subsequent figure-ground computation; the previous studies did not observe responses to ambiguous figure-ground stimuli which contained figure-ground cues but no true figure-ground percept.

T-junctions are important cues for determining the figure-ground configuration of visual stimuli. They are created by the occlusion of one object by another and serve as potent cues for segmenting the scene into depth-ordered surfaces (Figure 2). Figure 2A shows a figure-ground stimulus which contains T-junctions (denoted by the red "T"s). The configuration of the T-junctions helps yield the interpretation that the dark rectangle is in front of the white rectangle. However, the stimulus in Figure 2B contains T-junctions, but its figure-ground interpretation is ambiguous. Therefore, the presence of figure-ground cues does not always yield an unambiguous figure-ground interpretation. Several questions arise: Would a cell respond differently to the stimulus in Figure 2A than to the one in Figure 2B if, in each case, its classical receptive field straddled the border between the two rectangles? What if the classical receptive field were located within the interior of one of these rectangles (e.g., the dark one)? Would the cell respond preferentially to the figure rectangle in Figure 2A, or would it respond as well to the dark rectangle of 2B with its surrounding figure-ground cues?

We began our investigation of the neuronal basis of figure-ground discrimination in experiments on a single, awake, fixating monkey. We recorded from single cells in areas V1 and V2, asking whether cells in these areas exhibited border ownership selectivity in unambiguous figure-ground stimuli (which had clear figure-ground percepts). We also asked whether cells exhibited true figure vs. ground selectivity,

FIGURE 2

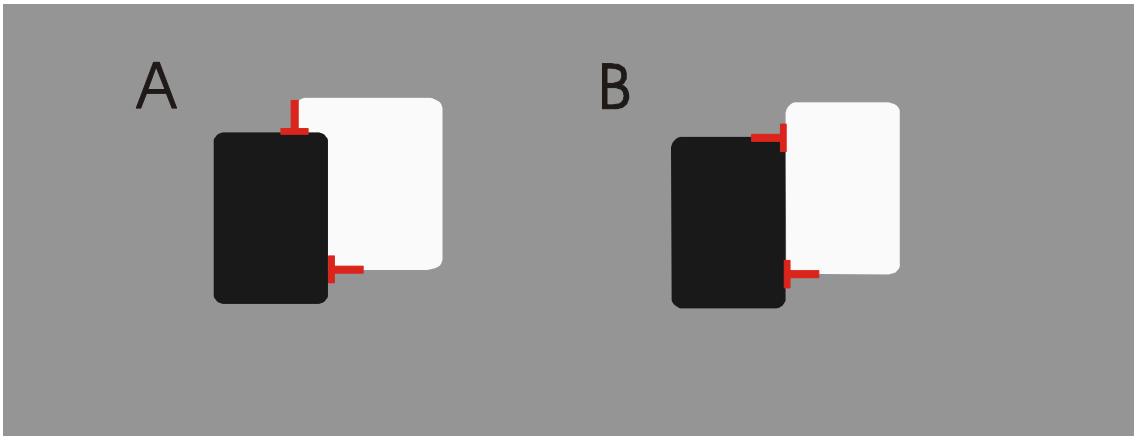

Figure 2: T-junctions. T-junctions are salient occlusion cues, and helpful in determining figure-ground aspects of the visual scene. However, T-junctions do not necessarily connote figure-ground. A: A stimulus which has clear figure-ground configuration, with T-junctions as indicated by the red "T"s. B: A stimulus which has no clear figure-ground configuration, yet still contains T-junctions.

responding differently when their receptive fields were located within the interior of figure vs. ground regions in these same stimuli. Furthermore, we also wished to control for possible responses to the presence of figure-ground cues (T-junctions) located outside the classical receptive field, even if the stimuli had no definitive figure-ground configuration. Therefore, we also tested cells when their receptive fields were positioned at the border or within the interior of ambiguous figure-ground stimuli, similar to the one shown in Figure 2B.

After the pilot experiment, we pursued the study more rigorously. We trained two monkeys to perform a behavioral match-to-sample task to ensure attention to the spatial aspects of the visual stimuli. We recorded spikes from cells in both V1 and V2. Again, we used stimuli with T-junctions and clear figure-ground configurations (unambiguous figure-ground stimuli), as well as stimuli with T-junctions, but no definitive figure-ground configuration (ambiguous figure-ground stimuli). The stimuli were positioned so that the cells' classical receptive fields were a) at the borders of the unambiguous figure-ground stimuli to determine whether the cells were selective for border ownership, or b) located entirely within the interior regions of the rectangles to determine whether the cells responded preferentially to figure regions vs. ground regions. Similarly, we recorded from the same cells when their receptive fields were positioned at the borders and within the interiors of the rectangles of ambiguous figure-ground stimuli. By testing the cells with both unambiguous and ambiguous figure-ground stimuli, we hoped to determine whether the cells were truly responding to the figure-ground configuration of the stimuli.

CHAPTER 2

Methods:

The majority of the results were obtained from two awake, behaving animals performing a match-to-sample task, and we begin this section by describing the methods for these experiments. The methods for the preliminary experiments on a single, awake, fixating animal are presented at the end of the section, emphasizing points of difference.

Match-to-Sample Experiment

Animal Subjects: Two adult, male rhesus monkeys (Macaca mulatta), weighing 10-11 kg, were used. Experimental protocols were approved by the Salk Institute Animal Care and Use Committee, and conform to the US Department of Agriculture regulations and to the National Institutes of Health guidelines for the humane care and use of laboratory animals.

Surgical Preparation: Procedures for surgery and wound care have been described in detail previously (Dobkins and Albright, 1994). To summarize, a head post and recording cylinder were affixed to the skull using stainless steel rails and screws and dental acrylic. Cranial magnetic resonance imaging (MRI) scans performed before surgery aided positioning of the recording chambers above areas V1 and V2. In animal 1, the chamber was over the left hemisphere and in animal 2, the chamber was located over the right hemisphere. A search coil for measuring eye position was surgically implanted in one eye in each animal using the method of Judge et al. (1980). The wire

leads were connected to a two-pin miniconnector which was affixed to the cranial implant with dental acrylic. After recovery from surgery, a craniotomy was performed to allow for electrode passage into areas V1 and V2. All surgical procedures were performed under sterile conditions, and animals were given prophylactic antibiotics (30mg/kg Keflin during surgery at 2 hr intervals) and post-surgical analgesics (buprenophorine, 0.03 mg/kg, i.m., every 12 hr for 3 d).

Apparatus for Visual Stimulation and Electrical Recording: Visual stimuli were generated using the two-computer version of Cortex 5.93 (developed in the Laboratory of Neuropsychology, National Institute of Mental Health). The display monitor was a 21" Sony Multiscan 500PS, with a refresh rate of 75 Hz. It was set at 1600 X 1200 pixel resolution, with 50 pixels per degree of visual angle. It was calibrated using a Photo Research spectrophotometer (SpectraColorimeter PS650). Cortex 5.93 was used for behavioral control of the monkeys and for data acquisition.

Neural responses were recorded using tungsten electrodes with vinyl resin insulation, or platinum-iridium electrodes with glass insulation at the tip and vinyl resin elsewhere (Frederick Haer and Co, 100 mm, 3 megohm, 250 micrometer shank diameter, standard medium final taper angle) inserted through the dura by a hydraulic microdrive (Crist Instruments) and micropositioner (Kopf, model 650). The signal was amplified with an AC differential amplifier (Bak Electronics), and filtered (Krohn-Hite, model 3700, low freq cutoff 700 Hz, high freq cutoff 11500 Hz, and Quest Scientific Hum Bug, 50/60 Hz Noise Eliminator). Spikes were then sorted (Alpha Omega Multi Spike Detector), and processed (Alpha Omega Multi Channel Processor, Cortex 5.93) to

identify the spikes generated by particular cells. The algorithm used to sort the spikes was based on that developed by Worgotter et al. (1986). It compared the electrode signal continuously against a template, and reported a spike whenever a match between the signal and the template occurred. Single units were defined by a minimal error between the signal and the template based on a sum of squares difference. Multi unit activity (firing from 2 or more neurons) caused greater error between signal and template.

MATLAB (The MathWorks, Inc.) software was used for all data analysis.

Eye Position: The eye coil system was calibrated at the beginning of each recording session, and eye position was monitored at all times using the scleral search coil technique, using dual power oscillators and phase sensitive detectors (CNC Engineering). Search coils were implanted in one eye of each animal. During data collection, the monkey was required to fixate within a window whose diameter was 0.8 degrees of visual angle or smaller—some recordings were taken with 0.75 and 0.7 degree diameter windows. Fixation within this window was required during the mapping of the receptive field, as well as the behavioral task. Trials were aborted without reward if the animal broke fixation at any time, or if the animal made a micro saccade >0.3 degrees within the fixation window.

Receptive Field Location and Cell Sampling in V1 and V2: All receptive fields were at eccentricities < 2 degrees in the contralateral visual field (right for animal 1, left for animal 2). Cells were methodically sampled across a 4 mm X 4 mm area in the chamber of each animal. Receptive field maps from each animal were reconstructed to determine

the boundaries between V1 and V2. The transition from V1 to V2 was characterized by the progression of receptive field location first towards the midline and then away from it as the electrode moved posteriorly to anteriorly. The transition was also marked by a sudden and marked increase in receptive field size.

Receptive Field Mapping: The approximate spatial location and preferred orientation of the receptive field was first determined by hand mapping using a white bar stimulus (65 cd/m2) on a grey background (10 cd/m2). The receptive field was then more carefully mapped using an automated computer-controlled stimulus that worked as follows: The background was grey, as before, and with the same luminance. A 2 X 2 degree invisible grid, significantly larger than the receptive field, was superimposed on the receptive field of the neuron. The grid was divided, like a checkerboard, into 81 compartments (9 X 9). A light or dark square stimulus (0.22 X 0.22 degrees) was presented at a random location on the grid for 100 msec, followed by another 100 msec with no stimulus, only a grey background. This process of stimulus, no stimulus was repeated until all 81 locations on the checkerboard grid had been sampled by both light and dark square stimuli. An additional 18 "blank" trials, where no square stimulus appeared but the background remained the same grey, were added randomly throughout the sampling session, to obtain the baseline firing rate of the neuron. In total, this entire process was repeated 3 times, so that each location on the checkerboard was sampled 6 times—3 times by white squares, and 3 times by black squares. The entire mapping procedure was then repeated, with a grid which measured 6 X 6 degrees. MATLAB software was then used to determine the spatial extent of the receptive field—by first plotting and then examining the neuron's

separate responses to the light squares and the dark squares (to determine on/off regions), and then to their difference. Two separate analyses were then performed. A difference of gaussians (DOG) model was fit to the neuron's profile of spatial responsiveness, and a separate thresholding paradigm was also used. In the DOG model, the DOGs could be oriented along any axis and could have any width along both the major and minor axes. Excitatory and inhibitory regions were determined by measuring the cell's responsiveness above or below its baseline firing level when a light or dark square was flashed at a particular spatial location. DOGs were then iteratively fit (through 2000 rounds) to the neuron's excitatory and inhibitory responses to the light squares as well as to the dark squares. The thresholding paradigm determined whether the cell's response at a given location on the grid was more than 2 standard deviations above (or below) the baseline firing rate of the neuron. In this case, this was considered a positive (or negative) response at that spatial location. From the results of the hand mapping as well as the computer analyses, the preferred orientation of the neuron was determined and matched to the closest among 0, 45, 90, 135 degrees. The neurons were not classified into categories such as "simple," "complex," or "hypercomplex". An example of a neuron's receptive field is shown in Figure 3. Because the monkey was fixating within a 0.8 degree window, the maps incorporate the jitter in the receptive field position due to small eye movements, which caused an apparent increase in the receptive field size. In Figure 3A, the 3 columns represent the actual data from a cell, the DOG fit, and thresholded data respectively. In Figure 3B, the results from the DOG fit are superimposed upon an outline of one of the stimuli (drawn to scale) to demonstrate that the T-junctions were not within the receptive field.

FIGURE 3

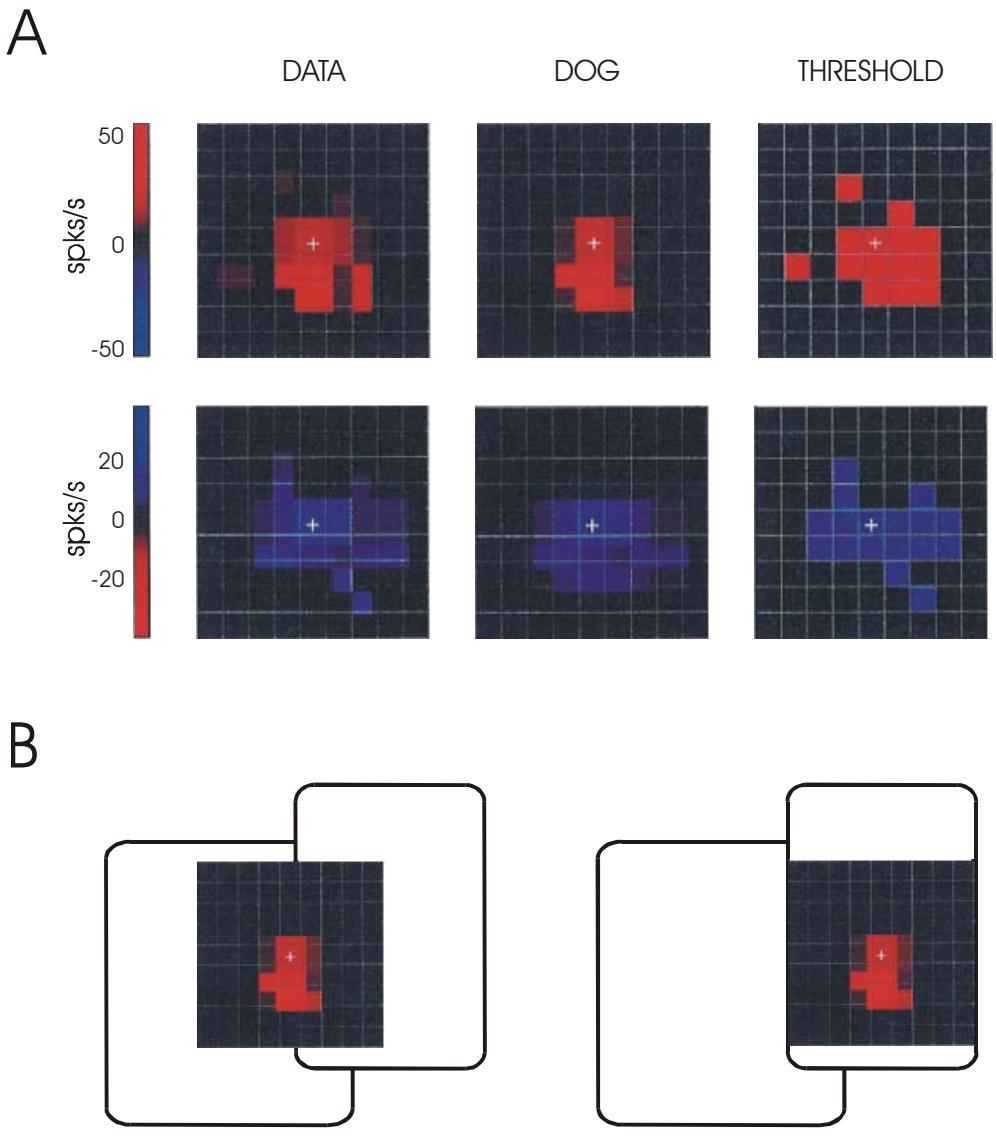

Figure 3: Receptive field maps of a cell. A: The first row contains responses of the cell to a 0.22 X 0.22 deg light square, which was randomly shown at all locations on a 9 X 9 grid. The color red signifies enhanced firing. The second row contains responses of the same cell to a 0.22 X 0.22 deg dark square similarly shown at all locations on the grid. In this case, the color blue signifies enhanced firing. The three columns contain the data itself, the DOG fit, and a thresholded version of the data respectively. B: The results of a DOG fit to the "on" responses are superimposed upon the outlines of actual stimuli (drawn to scale) as an example, demonstrating that the T-junctions are not located within the receptive field.

Visual Stimuli: A total of 8 visual stimuli were used in the match-to-sample experiment: 4 figure-ground stimuli (Figure 4AA), and 4 non-figure-ground stimuli (Figure 4BB). The figure-ground stimuli consisted of a pair of overlapping rectangles. In each pair, one rectangle was light, the other dark, one was in front (figure), and the other behind (ground). Therefore, each pair of rectangles had a luminance polarity (e.g., black on left, white on right) and a true figure-ground configuration (e.g., figure on left, ground on right). The non-figure-ground stimuli also consisted of a pair of rectangles with similar luminance properties (one light, and one dark), but with no clear figure-ground properties. After the neuron's preferred orientation was determined using a combination of the hand mapping and the computer mapping described above, stimuli were chosen to match the orientation of the border between the two rectangles to the preferred orientation of the neuron. The stimulus was then positioned at one of three locations in space so that the cell's receptive field lay either at the border between two rectangles or in the interior of one rectangle or the other.

In the figure-ground stimuli, the figure rectangles measured 2 by 3 degrees, significantly larger than the sizes of the receptive fields of the cells (0.1–0.8 degrees). The ground-rectangle also measured 2 by 3 degrees, but contained small extensions which created T-junctions causing the ground-rectangle to be occluded by the figure. The small extensions also made the total surface area of the ground rectangle appear somewhat larger than that of the figure rectangle.

Figure 5 shows the various arrangements of the T-junctions in the figure-ground and non-figure-ground stimuli. In this figure, 4 representative stimuli with the same luminance polarity (dark left, light right) are shown. A and C are figure-ground stimuli,

FIGURE 4

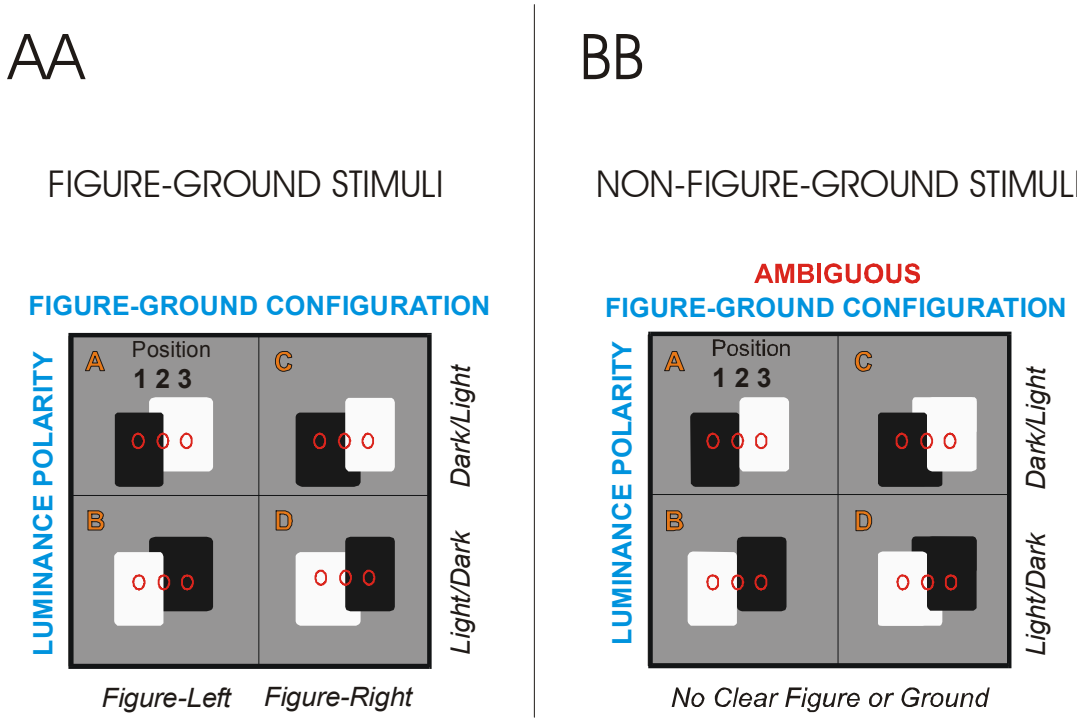

Figure 4: The stimuli used in the match-to-sample experiments. AA: Unambiguous figure-ground stimuli. These stimuli have clear figure-ground aspects; one rectangle is in front of the other. There are two luminance polarity configurations (e.g., black on left/white on right, white on left/black on right), and two figure-ground configurations (e.g., figure on left/ground on right, ground on left/figure on right). Note that these stimuli are presented at this orientation as a prototypical example; in the experiment the stimuli were oriented along the preferred orientation of the neuron. The red circles in the interiors of the stimuli represent receptive field locations. Three locations were used: one at the border between the two rectangles (position 2), and the other two in the interiors of the rectangles (positions 1 and 3). The size of the figure rectangle (the one in front) measured 2 deg X 3 deg, significantly larger than the size of the receptive fields. The size of the ground rectangle (the one behind) was the same as that of the figure rectangle, but with an additional small area extension. This extension created the occlusion effect by forming a T-junction, along with the illusion that the ground rectangle had a significantly greater surface area than the figure rectangle. For example, in A, the small extension of the white rectangle in back, located just above the upper right hand corner of the black rectangle, creates a T junction. This leads to the appearance that the black rectangle is in front of a larger white rectangle. BB: Ambiguous figure-ground stimuli. The figure-ground ordering of these stimuli is ambiguous; it is unclear which shape is in front of the other. Like the unambiguous figure-ground stimuli, these stimuli have two different luminance polarity configurations and two different spatial configurations.

FIGURE 5

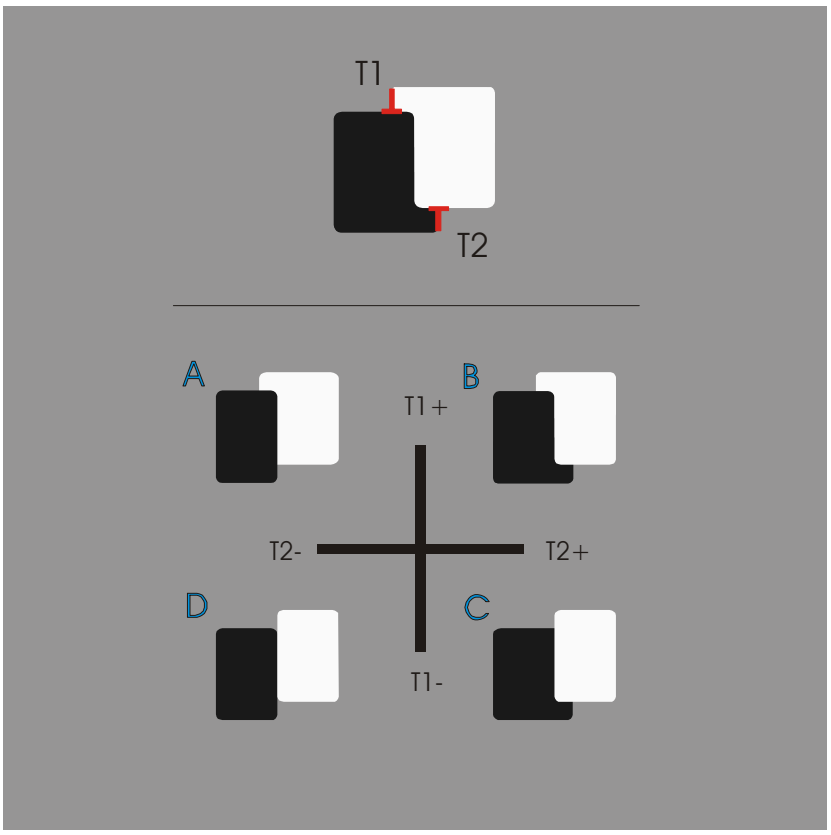

Figure 5: T-junction configurations in unambiguous and ambiguous figure-ground stimuli. *A, B, C, D* are representative unambiguous and ambiguous figure-ground stimuli with the same luminance polarity, dark-left, light-right. *A* and *C* are unambiguous figure-ground stimuli, and *B* and *D* are ambiguous figure-ground stimuli. *A* and *C* can be considered to be comprised of one figure rectangle and one ground rectangle, whereas *B* and *D* can be considered to be comprised of two ground rectangles and two figure rectangles respectively. The stimuli can also be described in terms of the two T-junctions, T1 and T2, as depicted in red in the example stimulus at the top of the figure. Going clockwise, *A* has T1 but not T2 (T1+,T2-). *B* has both T1 and T2 (T1+,T2+). *C* does not have T1 but does have T2 (T1-, T2+). Finally, *D* has neither T1 nor T2 (T1-,T2-). T1 and T2 are created by small area extensions to *D*. For instance, the difference between *D* and *A* is the small extension of the white rectangle above the top right corner of the black rectangle. The difference between *A* and *B* is the addition of a small extension of the black rectangle below the bottom left corner of the white rectangle. Similarly, *C* is the same as *B* minus the small extension of the white rectangle above the top right corner of the black rectangle. Finally, *D* is the same as *C*, minus the extension of the black rectangle below the left corner of the white rectangle.

and B and D are non-figure-ground stimuli. A and C can be considered to be comprised of one figure rectangle in front of one ground rectangle, whereas B and D, which have no readily apparent figure-ground ordering, can be considered to be comprised of two ground rectangles and two figure rectangles respectively. All four stimuli are related by the presence or the absence of two particular T-junctions, T1 and T2, shown in red, in the example stimulus at the bottom of the figure. Starting at A and going clockwise, we see A contains T1 but not T2. This is indicated in the space shown as (T1+, T2-). B contains T1 and T2 (T1+,T2+), C does not contains T1 but does contain T2 (T1-,T2+), and D contains neither T1 nor T2 (T1-,T2-). The presence of T1 and T2 are due to small area extensions off of D. For instance, A has the same area as D with a small additional area of white extending above the top right corner of the black rectangle. Similarly, stimulus B has the same area as A with an additional small area of black extending below the bottom left corner of the white rectangle. C has the same area as B minus the small white area above the top right corner of the black rectangle. Finally, D has the same area as C minus the small black area below the bottom left corner of the white rectangle. All the stimuli can be fundamentally derived from C with small area extensions. However, due to the strong appearance of occlusion in the figure-ground stimuli, it appears that the ground rectangle has a significantly larger surface than the figure rectangle.

Behavioral Task: Two male rhesus macaque monkeys were trained on the match-to-sample task, diagrammed in Figure 6. Throughout training and subsequent recording sessions, the monkey was seated in a standard primate chair (Crist Instruments), 57 cm away from the screen, with the head post rigidly fixed to the frame of the chair. At the

FIGURE 6

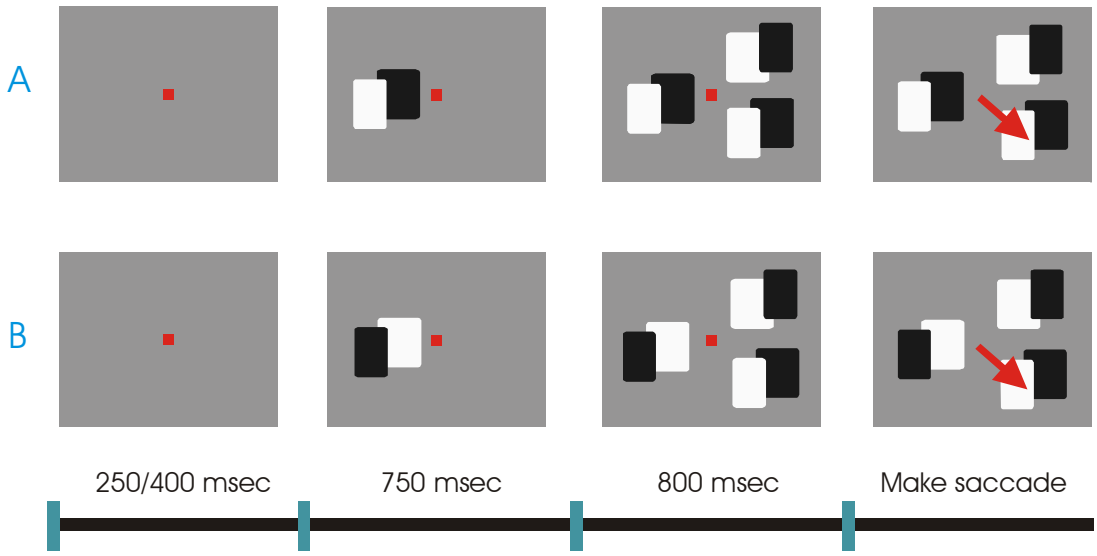

Figure 6: Match-to-sample paradigm. *A* depicts the match-to-sample paradigm when the sample, match, and distractor all had the same luminance polarity. *B* depicts the same paradigm when the match and distractor had the opposite luminance polarity as the sample. The general paradigm is as follows: A small, red fixation spot appeared. The monkey had 3 sec to achieve fixation, and had to hold fixation for 250 msec (earlier data) or 400 msec (later data). At this point the sample appeared, and remained on for 750 msecs, while the monkey maintained fixation. Then a match and a distractor came on. The match and distractor both had either the same luminance polarity as the sample (*A*) or the opposite luminance polarity (*B*). The match is defined as the stimulus which had the same spatial configuration as the sample, regardless of its luminance polarity. In the examples shown, the matches are the stimuli towards the bottom right of the figures in the third time frame. All three stimuli--sample, match, and distractor--remained on for an additional 800 msecs, and the monkey continued to fixate. The fixation spot then extinguished, and the monkey was required to make an eye movement to the match in order to receive a juice reward.

beginning of a trial, a small, red fixation spot (0.2 deg in diameter) appeared in the center of the screen. The monkey had three seconds to achieve fixation, and was required to hold its gaze on the spot for 250 msec (earlier data) or 400 msec (later data). At this time, a sample stimulus consisting of a black and a white rectangle appeared alone on the screen. The white and black rectangles were presented on a grey background. The luminances—white = 65 cd/m2, black = 1.5 cd/m2, grey = 10cd/m2—were chosen so that the Michaelson contrasts of the rectangles on the background were the same. 750 msecs later, while the monkey continued to fixate, a match stimulus and a distractor stimulus appeared such that three stimuli were simultaneously present on the screen. 50% of the time both the match stimulus and the distractor stimulus had the same luminance polarity as the sample stimulus, and 50% of the time both had the opposite luminance polarity. The matching stimulus was the one with the same spatial configuration, regardless of its luminance polarity. After an additional 800 msecs, the fixation point was extinguished, and the animal's task was to make a saccade to the matching stimulus. As soon as the animal made a choice to one target, the other stimulus target was extinguished, such that only the original sample stimulus and the chosen target stimulus remained on the screen. If the saccade was to the correct target, the monkey received a juice reward, and the stimulus and the match remained on the screen for an additional 600/1000 (animal 1/animal 2) msecs before they were extinguished. If the monkey chose the incorrect target, no juice was given, and the stimulus and the incorrect choice remained on the screen for 250/300 msec (animal 1/animal 2) before being extinguished. If the animal broke fixation at any time during the trial, the trial was aborted and all the stimuli on the screen were immediately extinguished. The standard

intertrial interval was 1500 msecs, unless the animal chose the incorrect target, in which case the intertrial interval was increased to 2500 msec. The animal performed 10 trials at each of the three receptive field locations (1 at the border between the two rectangles, 2 in the interiors of the rectangles) for each stimulus.

Data Analysis: MATLAB software was used for analyzing all data. Firing rates were calculated over a time window of 50–1550 msec after stimulus onset, for each of the 10 trials at each location. In order to determine whether the cells exhibited significant tuning to the contrast of the stimuli as well as their figure-ground configuration, a balanced 2-way ANOVA ($p<0.05$) was performed on the responses from each cell at each of the three receptive field locations, using luminance polarity and spatial configuration as the two comparative parameters. In addition, the 2-way ANOVA was performed grouping all the responses when the receptive fields were located at figure locations and comparing them to those from the ground locations regardless of whether the receptive field was at position 1 or 3. A similar analysis was performed using the responses to non-figure-ground stimuli. Monte Carlo analyses were performed to verify the results of all the initial ANOVAs. This involved randomly shuffling all the responses from a single cell on a trial-by-trial basis, performing an ANOVA, and repeating this 2000 times. The number of times a significant result was obtained was then counted. If the total count was fewer than 5% of 2000 (=100), the initial results were considered significant.

Mean firing rates over this 50-1550 msec time window were also calculated, for the 10 trials at each of the three receptive field positions, for both the figure-ground and non-figure-ground stimuli. These rates were used to obtain a modulation index (MI) for

each stimulus condition. The MI was defined as the sum of the mean firing rates at the figure locations, minus the sum of the mean firing rates at the ground locations, divided by the sum of the firing rates at figure locations and ground locations. For instance:

$$MI = [(A + B) – (C + D)]/(A + B + C + D)$$

where A and B are mean firing rates at figure locations, and C and D are firing rates at ground locations. Histograms of the MIs provided independent evidence on whether the total population of cells studies in a given area contained a subpopulation that responded selectively to figure-ground configuration or figure-ground cues.

Preliminary Experiments on Fixating Monkey

The methods used in the preliminary experiments on a single, fixating monkey are the same as those described above except for the following:

Animal Subject: A single, female rhesus monkey, weighing 13 kg was used.

Surgical Preparation: A recording chamber was positioned above areas V1 and V2 in the right hemisphere.

Eye Position: The monkey was required to fixate within a 1.0 degree diameter window.

Receptive Field Location: All receptive fields were at eccentricities < 4 degrees in the contralateral visual field. While the cells were sampled at various locations within the chamber, they were not methodically sampled to determine the boundaries between V1

and V2; the distinction between V1 and V2 cells was primarily based on the size of the receptive field vs. eccentricity, as well as the size of the receptive field vs. depth of electrode penetration according to Gattass et al. (1981). Cells located more superficially in the cortex and having smaller receptive field sizes (< 1 degree diameter) were considered to be in V1, whereas cells encountered deeper in the cortex, and having larger receptive field sizes (> 1 degree) were considered to be in V2. Because during the electrode penetrations, there were consistent periods of silence between the superficial and deep regions, as well as measurable and large increases in receptive field on moving from superficial to deep, we considered this adequate for assigning cells to V1 or V2 in these preliminary experiments.

Receptive Field Mapping: The approximate spatial location and preferred orientation of the receptive field were determined by hand mapping using a white bar stimulus (65 cd/m2) on a grey background (10 cd/m2).

Visual Stimuli: A total of 8 visual stimuli were used, as shown in Figure 7. While they were similar to the ones used in the match-to-sample task, they were not identical. The contours in these stimuli were rectilinear, rather than rounded, as in the stimuli for the awake-behaving monkeys. Also, two of the ambiguous figure-ground stimuli were somewhat different, as shown in Figure 7BB-C, and 7BB-D. In these stimuli, the rectangles are adjacent to each other. Compare these to the stimuli shown in Figures 4BB-C, and 4BB-D.

FIGURE 7

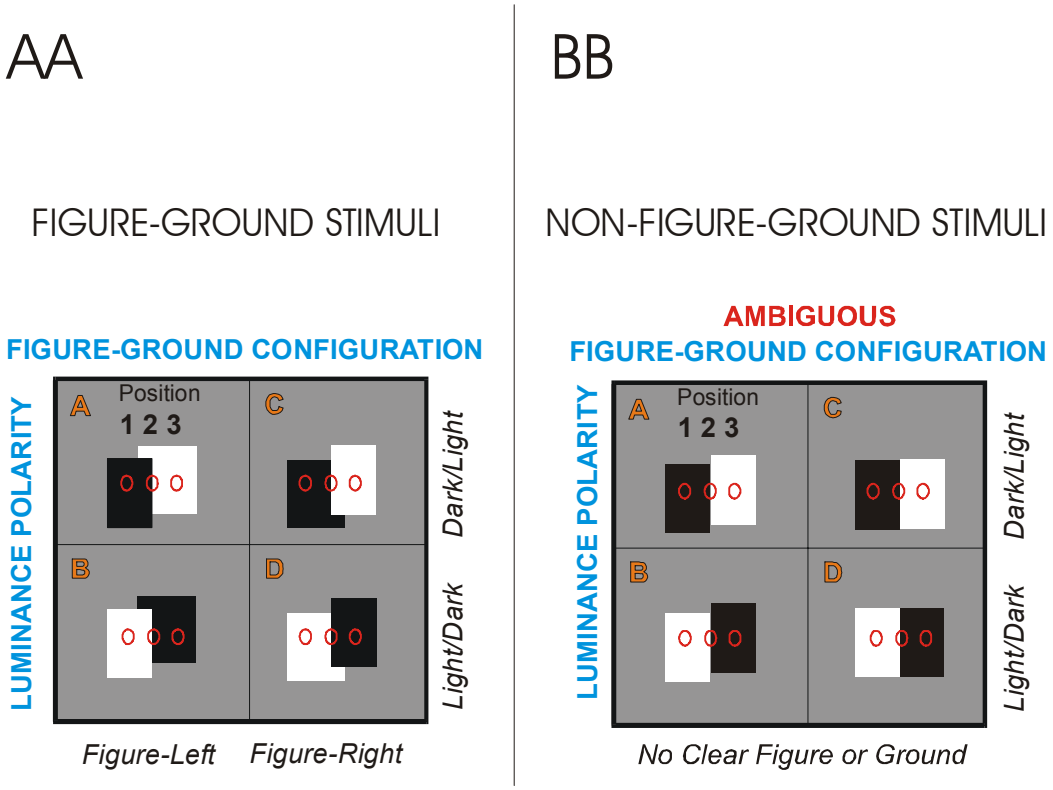

Figure 7: Visual stimuli used in the preliminary experiments on the fixating monkey. AA: The unambiguous figure-ground stimuli. These are exactly the same as the ones used for the match-to-sample experiment, but the contours are rectilinear. BB: The ambiguous figure-ground stimuli. The contours are also rectilinear, and the stimuli in C and D are somewhat different than the stimuli in the analogous stimuli in Figure 6. In this case, the rectangles are adjacent to one another.

Fixation Task: See Figure 8. The monkey was required only to fixate. At the beginning of the trial, a small, red fixation spot (0.2 deg in diameter) appeared, and the monkey had 3 seconds to achieve fixation. The monkey had to fixate on the spot for 500 msec, at which point a visual stimulus appeared. The monkey had to continue fixating on the spot for another 500 msec, while the stimulus remained on the screen. The fixation spot and the stimulus were then extinguished, and the monkey received a small juice reward.

Data Analysis: Mean firing rates were calculated over the time period of 50-500 msec after stimulus onset, for each of the 10 trials at each location. ANOVAs and modulation indices were calculated as in the match-to-sample experiments, but no Monte Carlo analyses were performed.

FIGURE 8

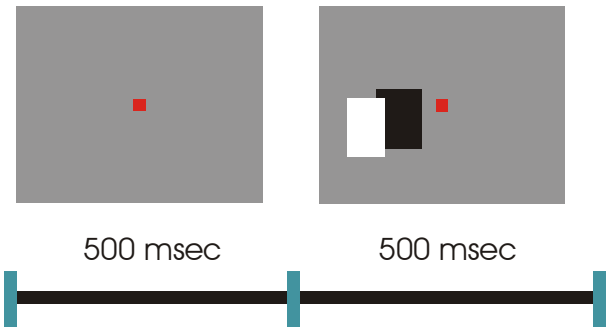

Figure 8: Fixation paradigm. A small fixation spot appeared, and the monkey had 3 sec to achieve fixation. The monkey had to hold fixation for 500 msec, at which point the visual stimulus appeared. The stimulus remained on for 500 msec while the monkey maintained fixation. The fixation spot and the sample were then extinguished. The monkey had to maintain fixation for the entire trial in order to receive a juice reward. The time interval between trials was 500 msec.

CHAPTER 3

Results:

Our goal was to comprehensively examine the possible role of neurons on V1 and V2 in figure-ground representation. We wished to determine whether cells in these areas exhibited "border ownership,"in which a neuron's response encodes to which side of a border between two rectangles the figure rectangle lay. We tested for border ownership when a cell's classical receptive field straddled a small region of the border between rectangles. We also wished to determine whether a neuron's responses encode which of two rectangles is the figure rectangle, when its classical receptive field is located entirely within one rectangle or the other. We began our studies recording from single neurons in areas V1 and V2 in an awake, fixating animal.

Results of Preliminary Experiments on a Fixating Monkey

Of the 20 V1 cells tested with unambiguous figure-ground-stimuli, one (5%) demonstrated border-ownership selectivity. With the receptive field at interior position 1, no cell showed figure-ground selectivity, and with the receptive field at interior position 3, 3/20 cells demonstrated figure-ground selectivity, for an average of 8%. Of the 51 V2 cells tested with unambiguous figure-ground stimuli, 10% (5/51) demonstrated border-ownership selectivity. An example of such a cell is shown in Figure 9. With its receptive field straddling the border between rectangles, this cell had enhanced responses when the figure was on the left, regardless of the luminance polarity of the two rectangles. Furthermore, 13% of the cells (5/51 at position 1 and 8/51 at position 3) showed figure-

FIGURE 9

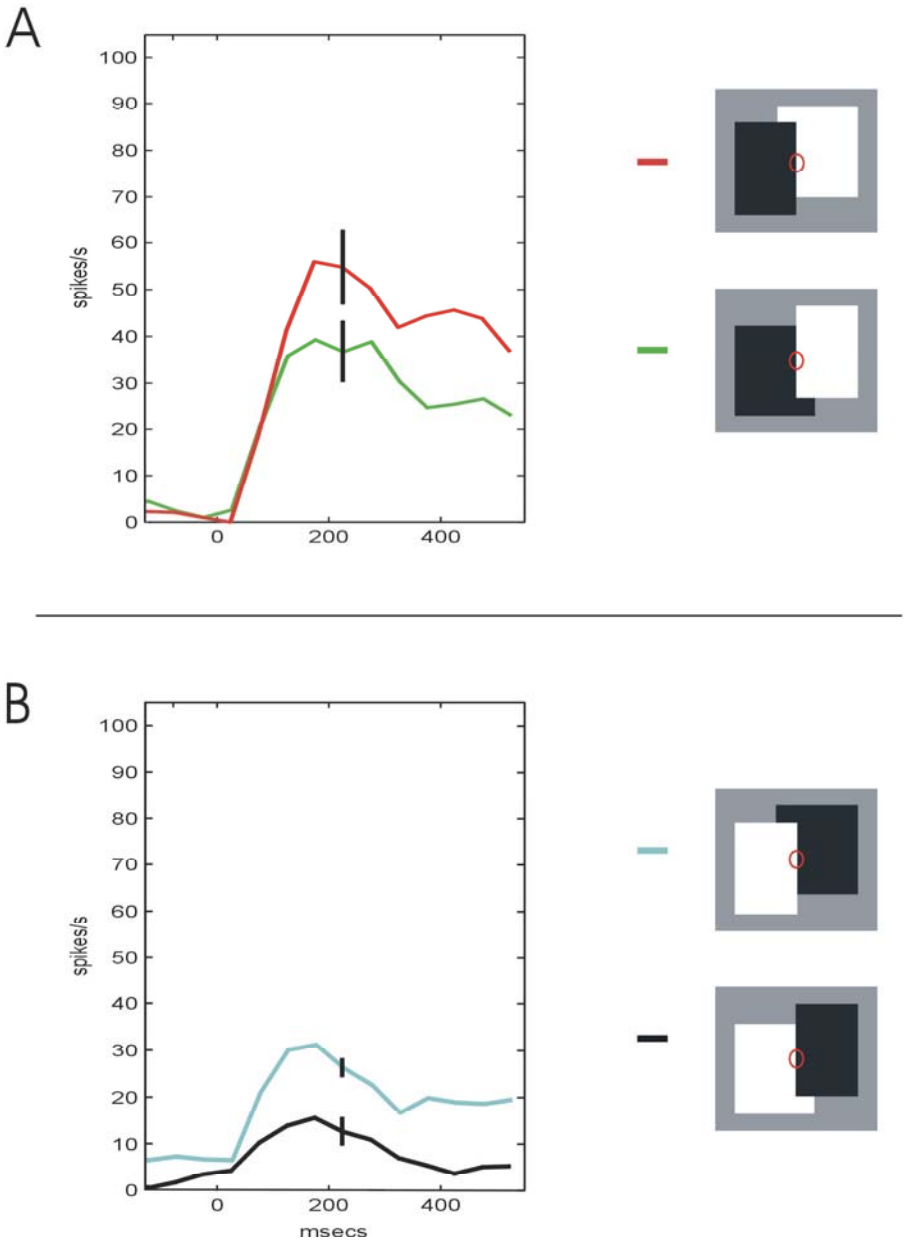

Figure 9: Responses from a V2 cell whose receptive field was at the border between rectangles in figure-ground stimuli. This cell was recorded from an awake, fixating monkey. A: The local image properties within the classical receptive field are identical, but the cell appears to prefer figure to the left. B: The same cell prefers figure to the left despite a switch in the rectangles' luminance. The stimuli here are identical to the stimuli in A, except for the change in luminance polarity. Each trace is averaged from 10 individual trails, Error bars show the standard error of the mean.

ground selectivity (ANOVA, $p<0.05$) with the receptive field in interior regions, with the majority of these cells (70%) preferring the interior of the figure region over the ground region.

We also tested the same cells in V1 and V2 with ambiguous figure-ground stimuli. Of the 20 V1 cells, again only 1 demonstrated apparent border-ownership, and none demonstrated selectivity for one of the interior regions, again suggesting that V1 may not play a major role in figure-ground interpretation. Of the 51 V2 cells, only 1 cell demonstrated apparent border ownership with ambiguous figure ground stimuli, although 15% (4/51 at one interior position and 11/51 at the other) demonstrated selectivity for one of the interior regions. An example of a cell demonstrating selectivity for interior position 3 is shown in Figure 10.

Very few cells in V1 demonstrated any selectivity to borders or interior regions in either the unambiguous or the ambiguous figure-ground stimuli. These preliminary results suggested that V1 may not play a strong role in the interpretation of figure-ground, or in detecting figure-ground cues such as T-junctions. More cells in V2 demonstrated border-ownership and figure-ground selectivity when tested with unambiguous figure-ground stimuli, indicating that V2 may be involved in figure-ground interpretation. Furthermore, some cells in V2 showed apparent selectivity to the ambiguous figure-ground stimuli, suggesting the existence of cells which responded not to the figure-ground aspects of stimuli, but rather figure-ground cues, such as T-junctions.

However, none of our fixation experiments nor those of previous workers could ensure the animals' attention to the spatial configuration of the stimuli. Therefore, we

FIGURE 10

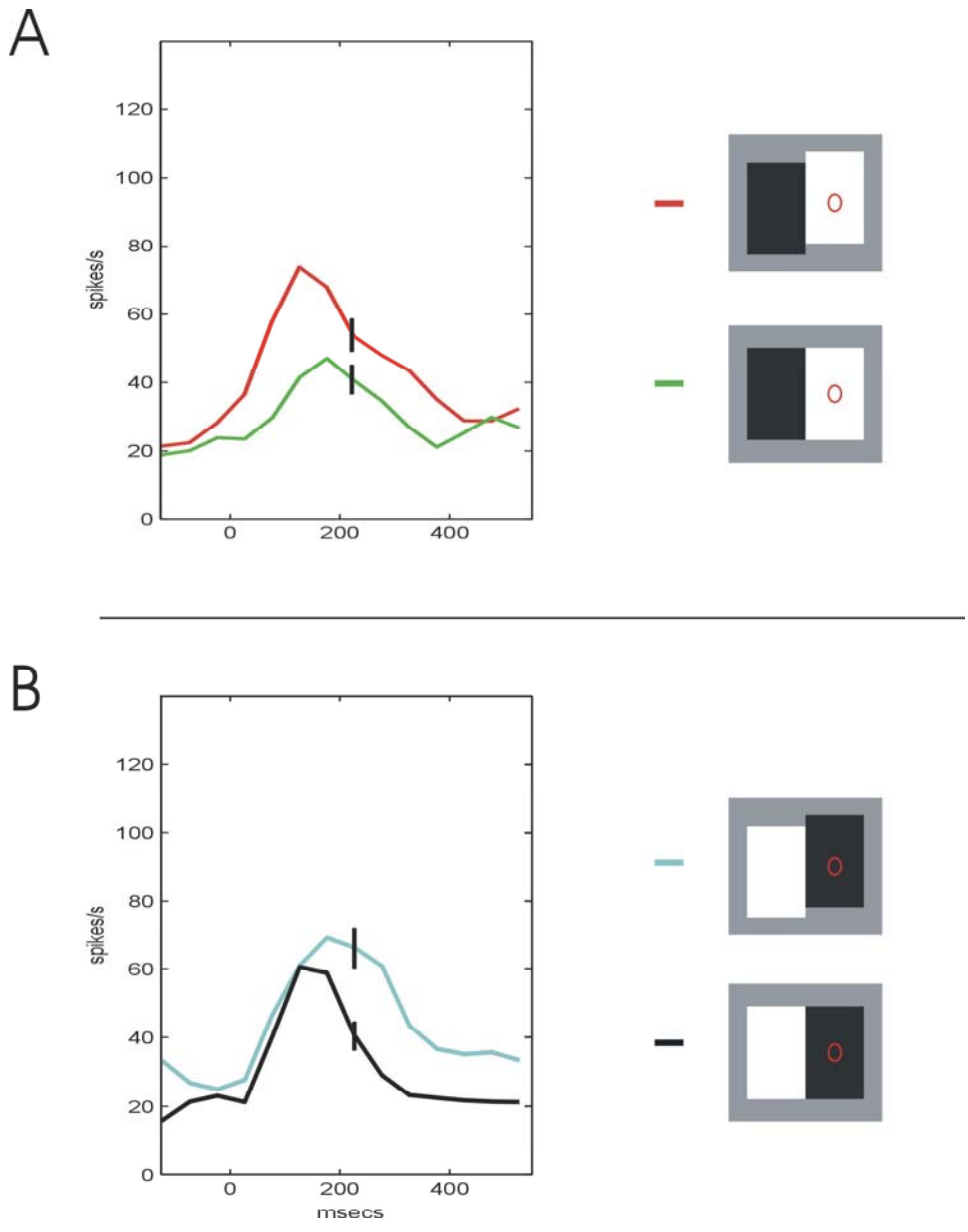

Figure 10: Responses from a V2 cell whose receptive field was in the interiors of rectangles in non-figure-ground stimuli. This cell was recorded from an awake, fixating monkey. A: This cell seemed to prefer the particular spatial configuration where the rectangles are offset rather than adjacent. B: The cell continued to prefer this spatial configuration, despite a change in luminance polarity. Traces were averaged over 10 trials.

pursued the study of figure-ground representation by training two additional monkeys to perform a behavioral match-to-sample task, which required the animal to discriminate between two very similar stimuli. The majority of our results, which were obtained from these behavioral experiments, are described in detail below.

Results from Match-to-Sample Experiments

We recorded from 103 sites in V1 and 170 sites in V2. Of these, 71 V1 sites were single-unit recordings, and 32 were multi-unit (2 or more cells) recordings. 128 V2 sites were single-unit and 42 were multi-unit. Because no significant differences were found between single- and multi-unit recordings, results from both types of recordings have been combined for the following analyses. The eccentricities of the receptive fields ranged from 0.8-1.7 degrees.

Tests for Border Ownership in V1

103 V1 sites were tested with unambiguous figure-ground stimuli, in which the borders between the black and white rectangles were centered within the CRF (Figure 4, position 2 on all stimuli). While the orientation of each stimulus was matched to the preferred orientation of the cell's receptive field along 0, 45, 90, and 135 degree axes (Cartesian coordinates), a prototypical stimulus aligned at 90 degrees will be shown hereafter for clarity.

Figure 11 shows the activity of a single cell in V1 whose receptive field was located at the border between rectangles in the unambiguous figure-ground stimuli. In Figure 11A, the cell's responses to a dark-light border are shown. The bar graph on the

left depicts the mean firing rates of the cell for each of the two stimulus conditions, and the tracings on the right show the cell's averaged responses over time. This cell's response was significantly stronger ($p < 0.05$ 2-way ANOVA, Monte Carlo analyses) when the figure was to the left of the boundary. Figure 11B shows the same cell's responses to a light-dark boundary. The cell's responses were again significantly stronger when the figure lay to the left of the boundary. The averaged sum of the responses from A and B are shown in C. This cell preferred figure to the left of the boundary, regardless of the luminance polarity of the stimulus.

Of the 103 sites which were tested with unambiguous figure-ground stimuli, 97 were also tested with ambiguous figure-ground stimuli. The ambiguous figure-ground stimuli had no readily apparent figure-ground arrangement. One set of stimuli appeared to be comprised of two figure rectangles, and the other set of two ground rectangles. Figure 12 shows the responses of a single unit to the borders of these stimuli. Figure 12A shows the cell's responses to a dark-light border between the two ambiguous figure-ground stimuli. As in Figure 11A, the bar graph on the left depicts the mean firing rates of the cell for the two conditions. The tracing on the right shows the cell's averaged responses over time. At this dark-light border, the cell demonstrated a significant preference ($p < 0.05$ 2-way ANOVA, Monte Carlo analyses) to the border between the two ground rectangles over the border between the two figure rectangles. Figure 12B shows the cell's responses to a light-dark border, and again this cell showed a significant preference for the border between the two ground rectangles. Figure 12C shows the averaged sum of the cell's responses in A and B. Overall, this cell demonstrated a significant preference for the borders of one set of ambiguous figure-ground stimuli,

FIGURE 11

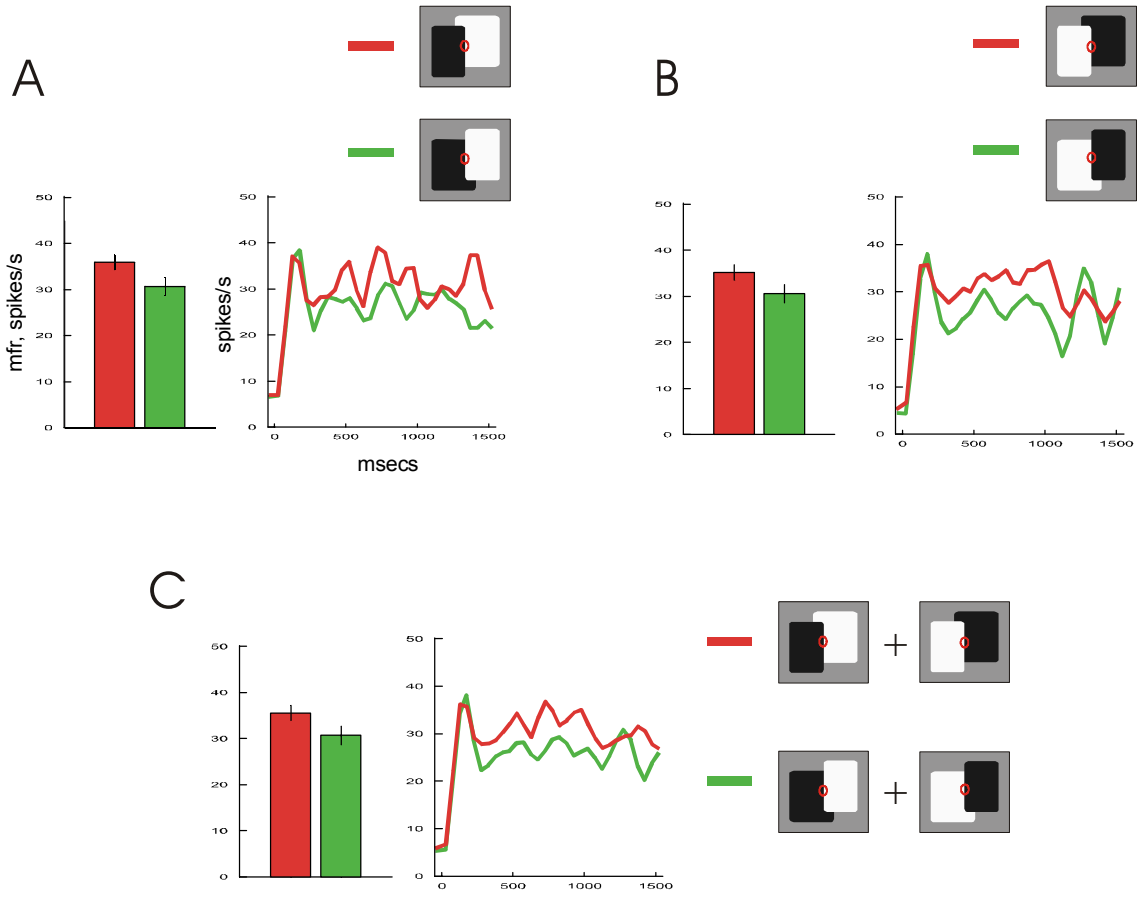

Figure 11: Responses from a V1 cell that demonstrated significant border-ownership selectivity in unambiguous figure-ground stimuli. In the graphs, the color red indicates those conditions in which the figure rectangle lay to the left of the border, and the color green indicates conditions in which the figure rectangle lay to the right of the border. A: The responses of the cell to a dark-light border. The bar plot on the left depicts the mean firing rate of the cell to two stimulus conditions over a 50-1550 msec time frame, averaged over 10 trials. The tracings on the right show the cell's averaged responses over time. This cell had a significantly greater response to the border when the figure rectangle lay to the left of the border. B: The responses of the cell to a light-dark border. Again, this cell had a significantly greater response to the border when the figure rectangle lay to the left. C: The averaged sum of the responses from A and B. This cell preferred the conditions in which the figure rectangle lay to the left of the border, regardless of the luminance polarity of the stimuli.

FIGURE 12

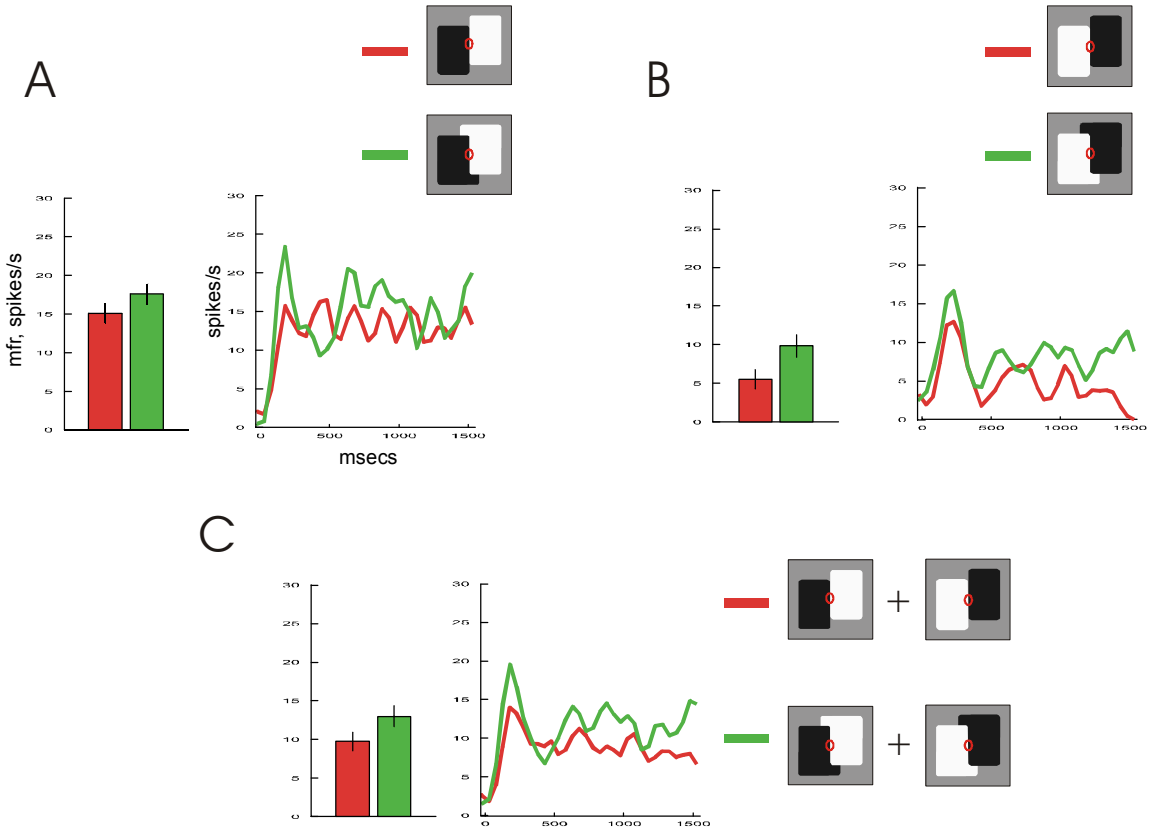

Figure 12: Responses from a different V1 cell demonstrating significant selectivity to borders in ambiguous figure-ground stimuli. Recall that the ambiguous figure-ground stimuli were comprised of either two figure rectangles or two ground rectangles. In the graphs, the color red indicates those conditions in which the stimuli consisted of two figure rectangles, and the color green indicates conditions in which the stimuli consisted of two ground rectangles. A: The responses of the cell to a dark-light border. As before, the bar plot on the left depicts the mean firing rate of the cell to two stimulus conditions over a 50-1550 msec time frame, averaged over 10 trials. The tracings on the right show the cell's averaged responses over time. This cell had a significantly greater response at the border between two ground rectangles as compared to the border between two figure rectangles. B: The responses of the cell to a light-dark border. Again, this cell had a significantly greater response at the border between two ground rectangles. C: The averaged sum of the responses from A and B. This cell preferred the borders between the two ground rectangles over the border between the two figure-rectangles regardless of luminance polarity. Because the figure-ground aspects of these stimuli are ambiguous, the cell may be selective for figure-ground cues located outside the classical receptive field rather than border ownership. The individual cells which responded to the borders of these ambiguous stimuli were largely different than those which responded to the borders of the unambiguous stimuli.

suggesting that this cell was responding to the figure-ground cues (T-junction configuration) rather than figure-ground border ownership.

Across the population of 103 sites tested with unambiguous figure-ground stimuli, 10% (10/103) of them showed significant border ownership selectivity irrespective of luminance polarity. As expected, the prevalence and the magnitude of luminance polarity selectivity far exceeded that for border ownership (as determined by the ANOVA and Monte Carlo analyses; see Methods)—51% (57/103) had a significant preference for one luminance polarity over the other, which is similar to the percentage (63%) obtained by Zhou et al. in a similar series of experiments. In addition, the ANOVA and Monte Carlo analyses revealed that 8% (8/103) of the cells had a significant interaction effect between border ownership selectivity and luminance polarity. This meant that these cells preferred a particular conjunction of border ownership and luminance polarity, such as black figure on the left, rather than figure on the left regardless of luminance.

In comparison, of the 97 sites which were tested with the ambiguous (control) stimuli, 6% (6/97) cells demonstrated a significant preference to the borders, regardless of luminance polarity. Again, the predominant effect found in this area was due to the difference in luminance polarities of the stimuli, with 54% (52/97 cells) having shown a significant preference for one luminance polarity over the other. 9% (9/97 cells) also showed a significant interaction effect between ambiguous figure-ground border selectivity and luminance polarity.

To quantify border ownership selectivity for unambiguous figure-ground stimuli across the V1 population, a modulation index was determined for each cell. The modulation index was the sum of mean responses for the conditions where figure was on

the left minus the sum of mean responses for the conditions where figure was on the right, divided by the sum of the responses for all these conditions. The relevant stimulus conditions are shown in Figure 13 on the left. A histogram of the modulation indices for all the V1 is shown in Figure 13 on the right. The histogram shows the indices for the cells demonstrating a significant effect through the ANOVA and Monte Carlo analyses (red), superimposed on the indices of all the cells (blue). If a preponderance of cells in V1 had border ownership preference, the entire distribution of modulation indices should be bi-modal, but instead the indices were normally distributed about zero (Jarque-Bera test of normality, $p = 0.5279$, t-test, $p = 0.8456$), with calculated mean = 0.0008, and median = 0.001. This indicated that the population as a whole was not selective for border ownership, although individual cells were selective.

The distribution of modulation indices for V1 cells tested with ambiguous figure-ground stimuli is shown in Figure 14. The indices across the entire population were again normally distributed about zero (Jarque-Bera test of normality, $p = 0.4036$, t-test, $p = 0.1065$), with mean = -0.0604 and median = 0.005 across the total population, indicating the population was not selective for borders, although again, individual cells did demonstrate selectivity.

In order to determine whether the cells were merely responding to the figure-ground cues rather than to the true figure-ground configuration of the stimuli, we compared the populations of cells which demonstrated significant effects at the borders of the unambiguous figure-ground and the ambiguous figure-ground stimuli. Presumably, if a cell were responding to certain aspects of the T-junction configuration, it might respond to both the unambiguous and the ambiguous figure-ground stimuli. However, the 10

FIGURE 13

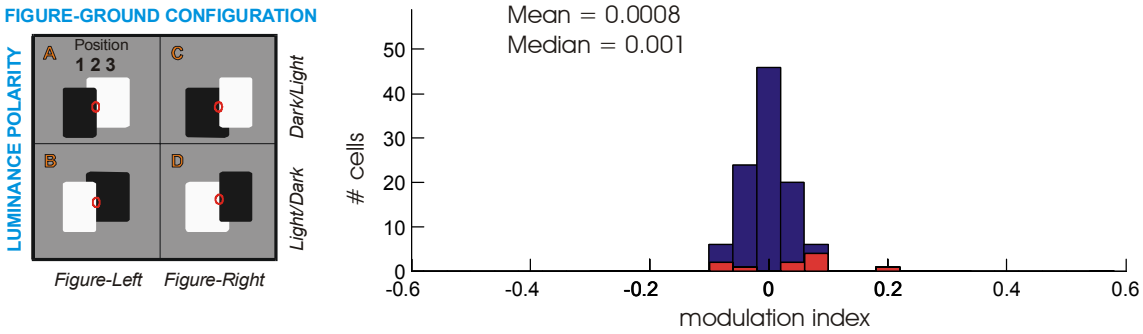

Figure 13: Modulation indices for all V1 cells tested with borders of unambiguous figure-ground stimuli. The modulation index is defined as: $[(a + b) - (c + d)]/(a + b + c + d)$ where a, b, c, and d represent the mean responses for the conditions depicted in A, B, C and D in the figure at the left. A histogram of the indices is shown on the right. The indices for all cells are shown in blue, and the indices for cells demonstrating border-ownership selectivity are shown in red. 10% of V1 cells demonstrated significant border ownership selectivity. However, the overall distribution was normally distributed and centered about zero, indicating the population as a whole was not selective for border ownership in the unambiguous figure-ground stimuli.

FIGURE 14

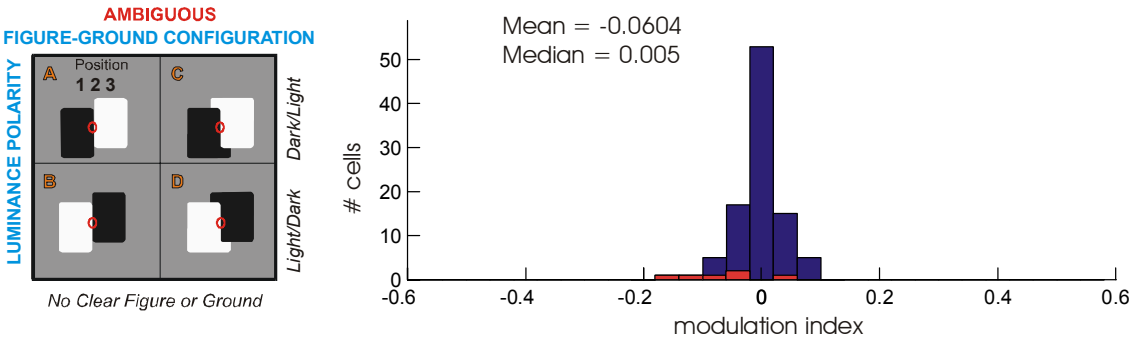

Figure 14: Modulation indices for all V1 cells tested with borders of ambiguous figure-ground stimuli. As previously, the modulation index was $[(a + b) - (c + d)]/(a + b + c + d)$ where a, b, c, and d represent the mean responses for the conditions depicted in A, B, C and D on the left of the figure, and a histogram of the indices is on the right. The indices for all cells are shown in blue, and the indices for cells demonstrating significant selectivity are shown in red. 6% of these cells demonstrated significant selectivity at the borders, indicating that the cells were responding to figure-ground cues, or T-junction configuration. However, the distribution was normally distributed and centered about zero, indicating that the population as a whole did not respond preferentially to the borders of ambiguous figure-ground stimuli.

cells which appeared to demonstrate border ownership were largely different from the 6 cells which showed an effect in the ambiguous figure-ground stimuli. Only one cell exhibited selectivity at the borders of both the unambiguous and ambiguous figure-ground stimuli. Because the cells with selectivity for the two different types of stimuli came from primarily from two independent populations, it appears that separate populations of cells may encode information about unambiguous and ambiguous figure-ground stimuli. Apparently, some cells demonstrate border ownership, and a separate class of cells responds to figure-ground cues (in this case, T-junction configuration). However, the number of cells demonstrating these kinds of effects was rather small across the population in V1. We wished to determine whether the effects would be similar in V2.

Tests for Border Ownership in V2

While we found in V1 particular cells which demonstrated border ownership selectivity at borders of unambiguous figure-ground stimuli, and other cells which demonstrated figure-ground cue (or T-junction configuration) selectivity at borders of ambiguous figure ground stimuli, the fractions of such cells were rather small across the population. To determine if differences between V1 and V2 existed, we pursued our study by recording from 170 cells in area V2. Of these cells, all were tested with unambiguous figure-ground stimuli, and 152 were also tested with ambiguous figure-ground stimuli.

Figure 15-1 shows an individual V2 neuron's responses to borders in unambiguous figure-ground stimuli. Figure 15-1A shows the cell's responses to a dark-

FIGURE 15-1

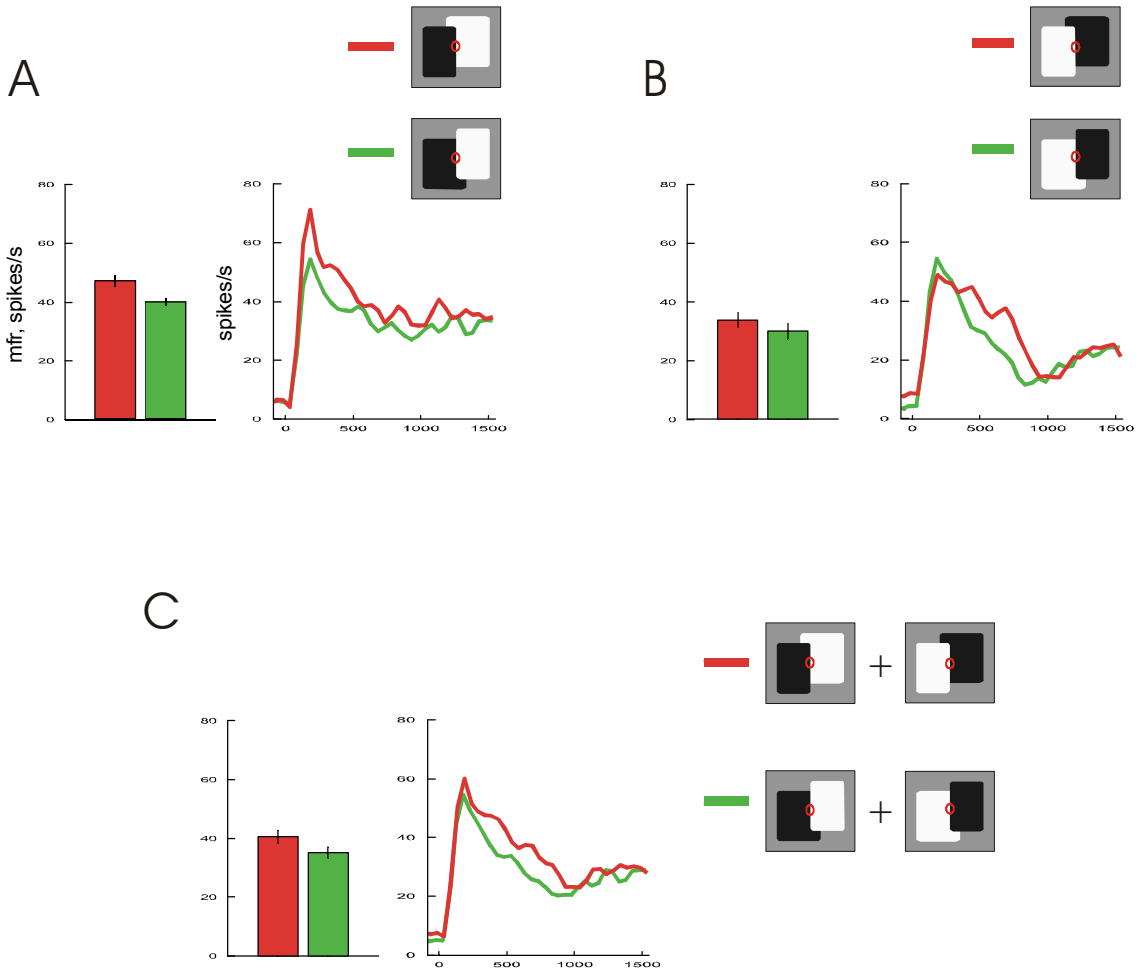

Figure 15-1: Responses from a single V2 cell demonstrating significant border-ownership selectivity in unambiguous figure-ground stimuli. As before, the color red indicates those conditions in which the figure rectangle lay to the left of the border, and the color green indicates conditions in which the figure rectangle lay to the right of the border. A: The responses of the cell to a dark-light border. This cell had a significantly greater response to the border when the figure rectangle lay to the left of the border than when the figure rectangle lay to the right of the border. B: The responses of the cell to a light-dark border. Again, this cell had a significantly greater response to the border when the figure rectangle lay to the left of the border. C: The averaged sum of the responses from A and B. This cell preferred the conditions in which the figure rectangle lay to the left of the border, regardless of the luminance polarity of the stimuli.

light border. This cell had a significant preference ($p< 0.05$ 2-way ANOVA, Monte Carlo analyses) for the conditions in which the figure was to the left of the dark-light border. In Figure 15-1B, the border was light-dark, but the cell again significantly preferred the condition in which the figure lay to the left of the border. The averaged sum of the responses in Figure 15-1A and Figure15-1B is shown in Figure 15-1C. This cell demonstrated significant border ownership selectivity, and preferred the situations in which the figure rectangle lay to the left of border, regardless of the luminance polarity of the stimuli. Figure 15-2 shows another example of a cell which demonstrated significant border ownership selectivity, but in this case, the cell preferred the stimuli in which the figure lay to the right of the border, regardless of the luminance polarity of the stimuli.

152/170 of the sites in V2 were additionally tested with ambiguous figure-ground stimuli to determine if there were V2 cells selective for figure-ground cues, or T-junction configuration, at borders rather than for border ownership. Figure 16 shows an example of a cell which exhibited a significant ($p<0.05$ 2-way ANOVA, Monte Carlo analyses) preference for one set of ambiguous figure-ground stimuli over the other. In Figure 16A, when the receptive field was located at dark-light borders, the neuron's response to the border between the two ground rectangles was significantly larger than its response to the border between the two figure rectangles. Similarly, this neuron preferred the border between the two ground rectangles at light-dark borders as shown in Figure 16B. The averaged sum of the responses obtained in Figs. 16A and 16B is shown in 16C. This neuron showed a significant preference for the border of one set of ambiguous figure-ground stimuli over the other, regardless of the luminance polarity. The neuron was apparently selectively responding to the figure-ground cues, or T-junction configurations,

FIGURE 15-2

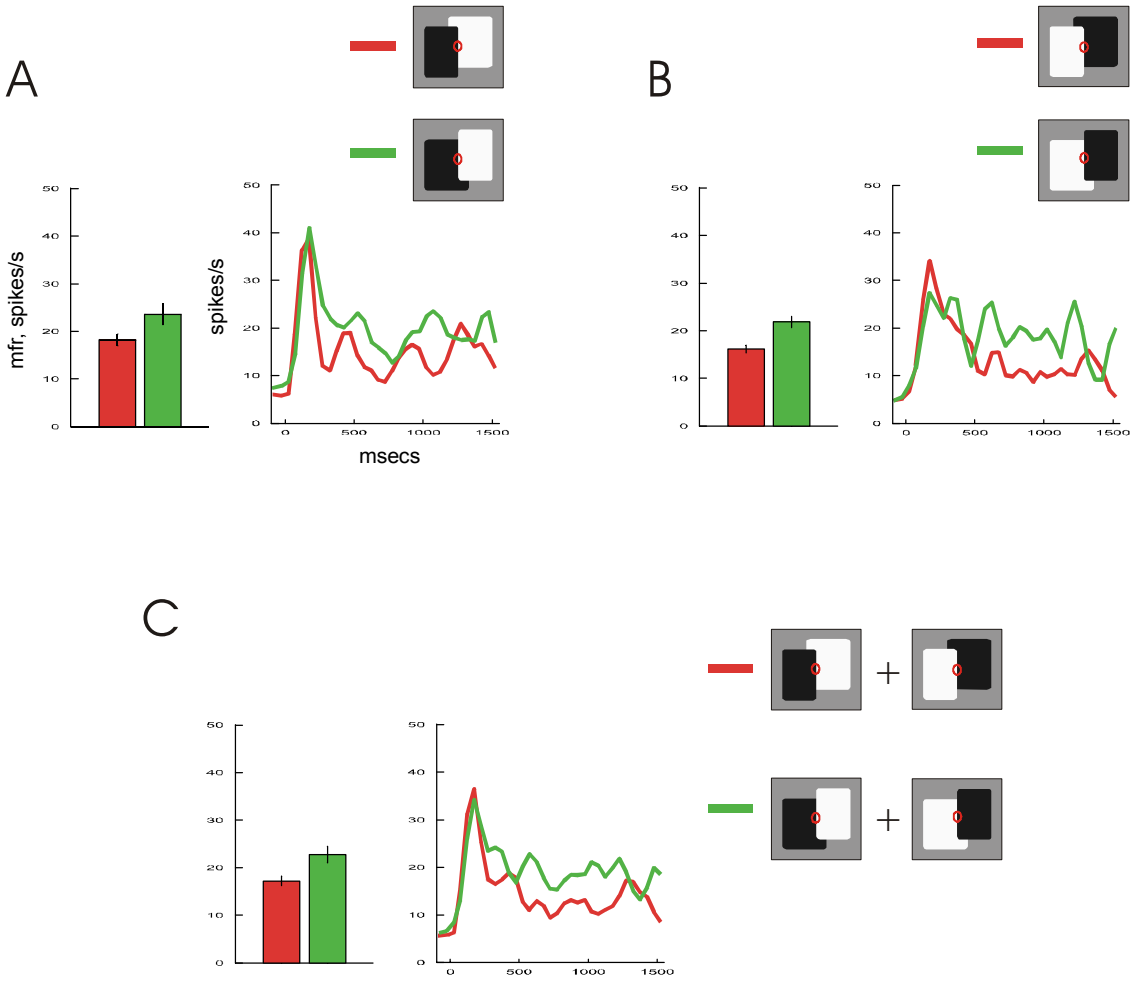

Figure 15-2: Behavior of another single V2 cell demonstrating significant border-ownership selectivity in unambiguous figure-ground stimuli. A: The responses of the cell to a dark-light border. This cell had a significantly greater response to the border when the figure rectangle lay to the right of the border than when the figure rectangle lay to the left of the border. B: The responses of the cell to a light-dark border. Again, this cell had a significantly greater response to the border when the figure rectangle lay to the right of the border. C: The averaged sum of the responses from A and B. This cell preferred the conditions in which the figure rectangle lay to the right of the border, regardless of the luminance polarity of the stimuli.

FIGURE 16

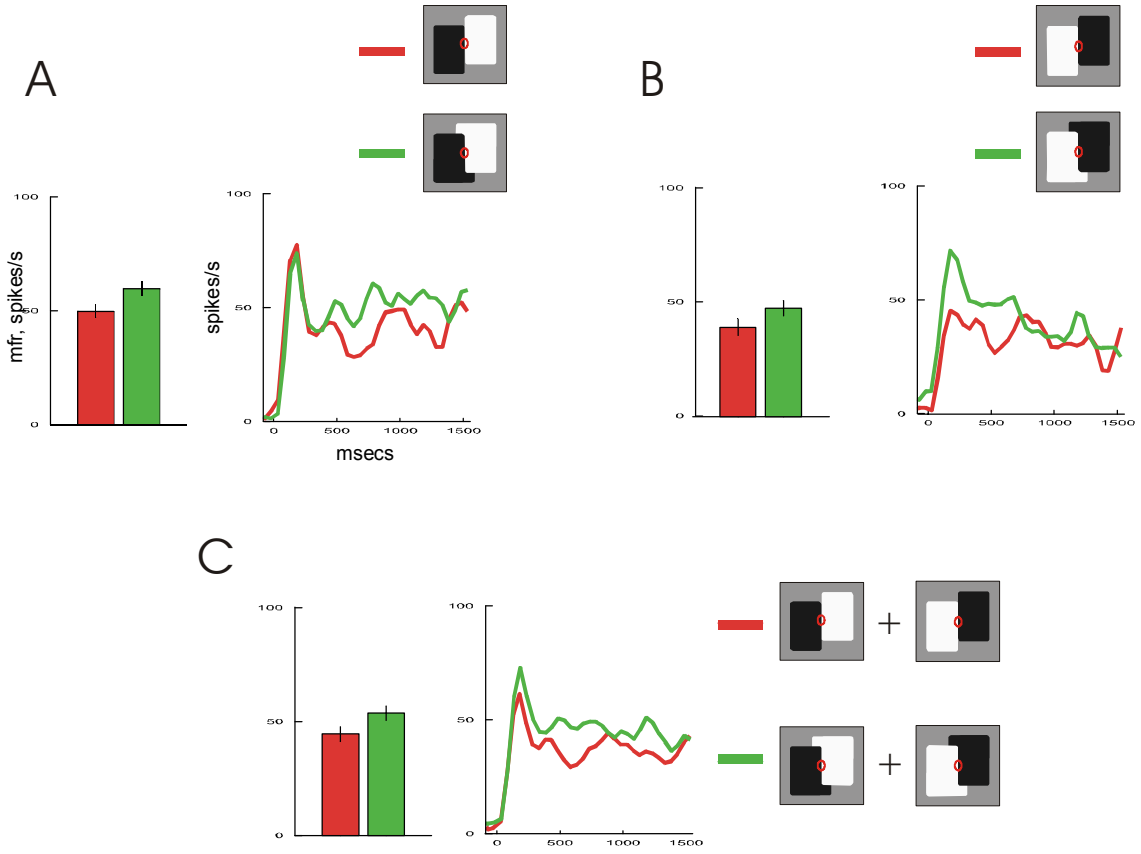

Figure 16: Responses from a single V2 cell demonstrating significant selectivity to borders in ambiguous figure-ground stimuli. A: This cell had a significantly greater response at the border between two ground rectangles as compared to the border between two figure rectangles. B: The responses of the cell to a light-dark border. Again, this cell had a significantly greater response at the border between two ground rectangles. C: The averaged sum of the responses from A and B. This cell preferred the borders between the two ground rectangles over the border between the two figure-rectangles regardless of luminance polarity. Because the figure-ground aspects of these stimuli are ambiguous, the cell may be selective for figure-ground cues located outside the classical receptive field rather than border ownership.

rather than border ownership, because these non-figure-ground stimuli lacked a definitive figure-ground configuration.

21% (36/170) of V2 cells demonstrated selectivity for border ownership, regardless of luminance polarity. This is about twice the percentage found in V1. However, as in V1, selectivity for luminance polarity was the dominant effect, with 48% (81/170 cells) responding preferentially to one polarity over the other. The 48% is comparable to the 66% found by Zhou et al. in V2. 17% (29/170 cells) showed selectivity for the interaction between figure-ground configuration and luminance polarity. These
cells demonstrated selectivity for particular combinations of figure-ground configuration and luminance.

11% (16/152 V2 cells) demonstrated significant preferences for borders of the ambiguous figure-ground stimuli, again approximately twice the percentage found in V1. The difference in luminance polarities was the dominant effect across the population, with 44% (67/152 cells) responding selectively to luminance polarity. Also, 7% (11/152 cells) showed an interaction effect between the ambiguous figure-ground borders and the luminance polarities, meaning these cells were selective for a particular combination of figure-ground configuration and luminance.

Modulation indices for the V2 cells tested with unambiguous figure-ground stimuli are shown in Figure 17. As in V1, the distribution of all the indices across the population was normal and centered at zero, with mean = 0.003 and median = 0.006 (Jarque-Bera test of normality $p = 0.2072$, t-test, $p = 0.9486$). Again, because the distribution of all the indices was not bimodal, this indicates that overall, the population

FIGURE 17

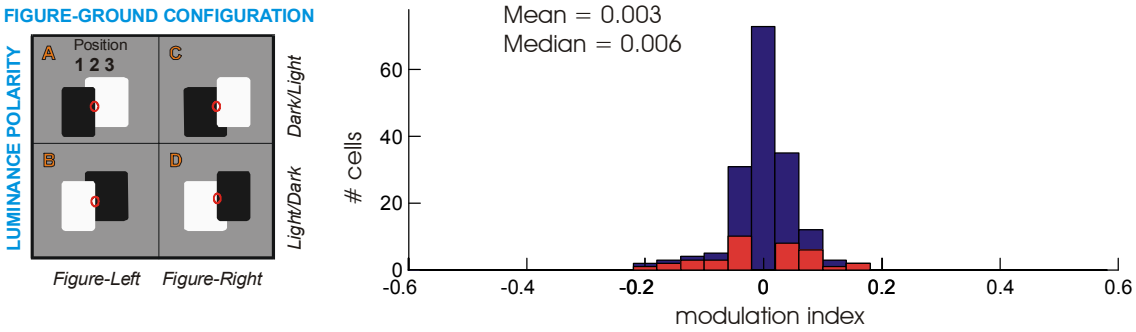

Figure 17: Modulation indices for all V2 cells tested with borders of unambiguous figure-ground stimuli. The indices for all cells are shown in blue, and for selective cells in red. 21% of these cells demonstrated significant border ownership selectivity, more than twice the percentage found in V1. However, the overall distribution was again normally distributed and centered about zero, indicating the population was not selective for border ownership in unambiguous figure-ground stimuli.

of V2 cells from which data was recorded did not demonstrate border ownership selectivity, although a fraction of cells within the population did.

Modulation indices for V2 cells tested with ambiguous figure-ground stimuli are shown in Figure 18. The distribution from all cells was normal and centered at zero, with mean = -0.0002 and median = -0.003 (Jarque-Bera test of normality $p = 0.1300$, t-test, $p = 0.9480$). As before, the population as a whole did not appear to be selective to the borders of the ambiguous figure-ground stimuli, even though a number of cells within that population demonstrated significant selectivity to these borders.

Were the cells which demonstrated selectivity to the borders of unambiguous figure-ground stimuli the same as those which demonstrated selectivity to the borders of the ambiguous figure-ground stimuli? Of the 36 cells which demonstrated selectivity to the borders of the unambiguous figure-ground stimuli, only 6 also demonstrated selectivity to the borders of the ambiguous figure-ground stimuli. Therefore, the populations which demonstrated selectivity for each type of stimuli were largely independent. This lends further evidence that separate populations of cells may be responsible for the determination of border ownership and for the detection of figure-ground cues which may lead to this percept.

V1 Responses to Interiors

103 V1 cells were tested with the interior regions of unambiguous figure-ground stimuli, and 97 of these were tested with the interior regions of ambiguous figure-ground stimuli.

FIGURE 18

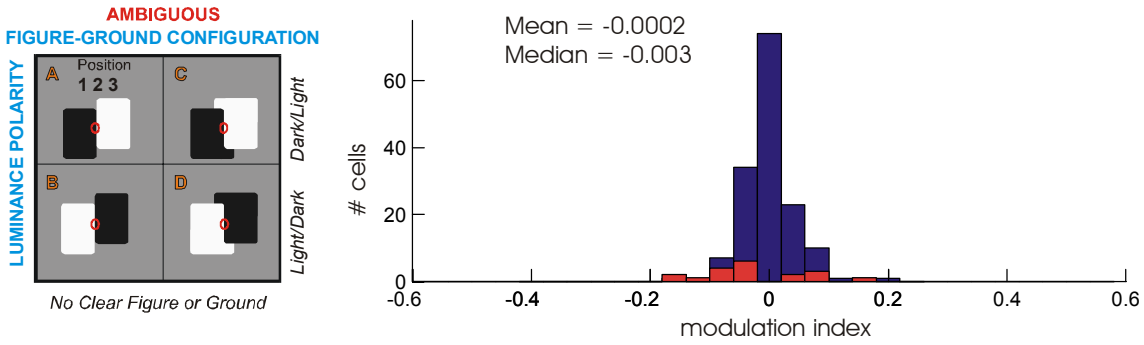

Figure 18: Modulation indices for all V2 cells tested with borders of ambiguous figure-ground stimuli. 11% of these cells demonstrated significant selectivity at the borders, nearly twice the percentage found in V1, indicating that the cells were responding to figure-ground cues, or T-junction configuration. However, the distribution was again normally distributed and centered about zero, indicating that the population as a whole did not respond preferentially to the borders of ambiguous figure-ground stimuli.

Overall, the responses to these stimuli were less vigorous because of the uniformity of the regions covering the classical receptive field. Nevertheless, the cells did respond to the interior regions of the rectangles in the stimuli. An example of a V1 cell's response to the interiors of unambiguous figure-ground stimuli is shown in Figure 19. This cell showed a significant preference ($p<0.05$, 2-way ANOVA, Monte Carlo analyses) for figural regions over ground regions at this spatial location in the stimuli (position 3). In Figure 19A, the cell's responses to light regions are shown. This cell responded preferentially when its receptive field was located in the figure rectangle as opposed to the ground rectangle. The same selectivity was apparent when the luminance polarity was switched, as shown in Figure 19B. Figure 19C shows the averaged sum of the responses from Figure 19A and 19B. This cell showed a significant preference for figure regions over the ground regions, regardless of the luminance of the stimuli.

97 of the 103 V1 cells were also tested with ambiguous figure-ground stimuli to examine whether they had figure-ground cue (or T-junction configuration) selectivity which might have caused the cells to respond preferentially to the interior regions of ambiguous figure-ground stimuli. However, the fraction of individual V1 cells that demonstrated significant selectivity to the figure-ground cues was smaller than that expected by chance. We therefore concluded that these cues did not lead V1 cells to respond preferentially to interior regions in the ambiguous figure-ground stimuli.

The statistics across the population of cells were as follows. There were two interior regions in each stimulus (at positions 1 and 3 in Fig 4). In the unambiguous stimuli, overall, only 2% (2/103) V1 cells demonstrated figure-ground selectivity at position 1, and 8% (8/103) demonstrated figure-ground selectivity at position 3. 5%

FIGURE 19

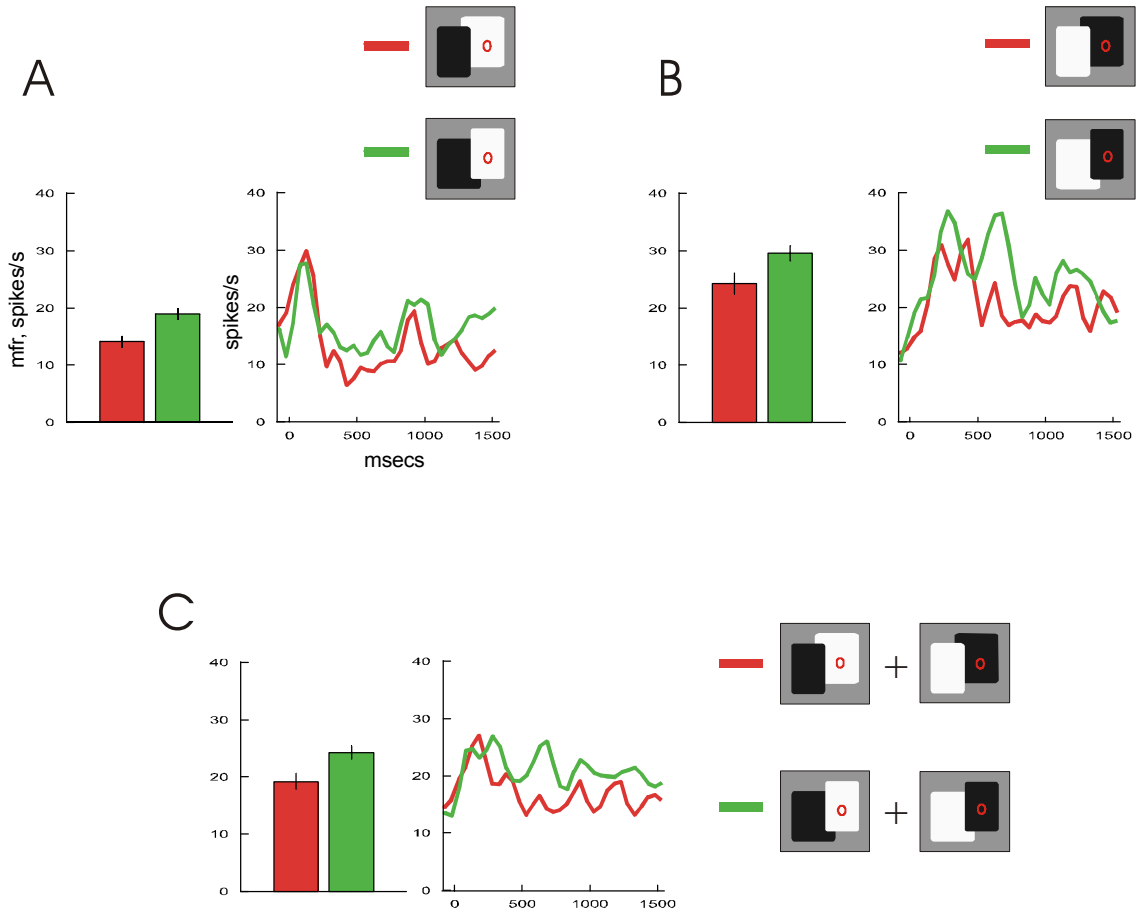

Figure 19: Responses from a single V1 cell demonstrating significant figure-ground selectivity in the interiors of unambiguous figure-ground stimuli. Red indicates conditions in which the figure rectangle lay to the left of the border, and the color green indicates conditions in which the figure rectangle lay to the right of the border. A: The responses of the cell to light interior regions at position 3. This cell had a significantly greater response when its receptive field was located in the interior of a figure region as compared to a ground region. B: The responses of the cell to dark interior regions at position 3. Again, this cell had a significantly greater response when its receptive field was located in the interior of the figure region. C: The averaged sum of the responses from A and B. This cell preferred the conditions in which its receptive field was located in figure regions over ground regions, regardless of luminance.

(5/103) showed figure-ground selectivity when the responses to all the figure regions were grouped and compared to the responses to all the ground regions, regardless of position. Although few cells showed figure-ground selectivity, it is notable that 80% (8/10) of cells which showed significant selectivity at positions 1 or 3 preferred the interior of figure regions over ground regions. Similarly, 80% (4/5) of cells which showed a figure-ground effect regardless of position preferred figure over ground. As expected, a large fraction of V2 neurons exhibited a luminance preference, with 81% (83/103) preferring one luminance over the other at position 1, and 80% (82/103) at position 3. 5% (5/103) demonstrated an interaction effect between the figure-ground and luminance aspects of the stimuli at each of the two positions.

Amongst the 97 V1 cells tested with ambiguous figure-ground stimuli only 3% (3/97) demonstrated significant selectivity to figure-ground cues (T-junction configuration) ($p<0.05$, 2-way ANOVA, Monte Carlo analyses) in the interior regions at position 1, and only 4% (4/97) did so at position 3. 2% (2/97) showed selectivity for the interior regions regardless of position, when the responses to all the figure rectangles were grouped and compared to the responses to all the ground rectangles (recall that the non-figure-ground stimuli consisted of either two figure rectangles or two ground rectangles). Because these percentages were smaller than the number expected by chance ($p = 0.05$), these "selectivities" were considered to be consistent with artifacts of the sampling. As with the figure-ground stimuli, a large fraction of the V1 cells responded preferentially to differing luminances. At position 1, 76% (74/97) of V1 cells preferred one luminance over the other, and at position 3, 76% (76/97) did. In addition, 7% (7/97) cells showed an interaction effect between the T-junction spatial configuration and

luminance at position 1, and 5% (5/97) did so at position 3. However, overall, because so few individual cells demonstrated significant selectivity to the configuration of the T-junctions, we concluded that these cues did not lead V1 cells to respond preferentially to interior regions in the ambiguous figure-ground stimuli.

Figure 20 shows modulation indices for the cells' responses to both interior regions. Figure 20A shows the indices for interior position 1, whereas Figure 20B shows the indices for interior position 3. If the population of these V1 neurons had an overall preference for figure over ground, the indices would be shifted towards positive values at position 1 and towards negative values at position 3. Or, if some neurons preferred ground over figure and others figure over ground, the populations might be bi-modally distributed. The histogram of position 1 indices shown in Figure 20A is statistically centered at zero (signed rank test, $p = 0.6475$), with mean = 0.004, and median = -0.001, indicating that at position 1, the cell population did not appear to demonstrate a significant figure-ground preference. Furthermore, only 2% (2/103) of individual cells appeared to show selectivity at this position. At position 3, however, the population appeared to have a slight preference for figure—the indices are normally distributed, (Jarque-Bera test of normality $p = 0.6675$), with mean less than zero (t-test, $p = 0.0038$). The calculated mean and median of this distribution were -0.01 and –0.008 respectively, but their both being very close to zero indicates that most cells in the population did not demonstrate appreciable selectivity. However, 7 of the 8 individual cells which showed selectivity at position 3 preferred figure over ground, lending some support to the results of previous studies which found V1 cells preferring figure regions. It is unclear why more cells seemed to prefer figure over ground in position 3 as compared to position 1.

FIGURE 20

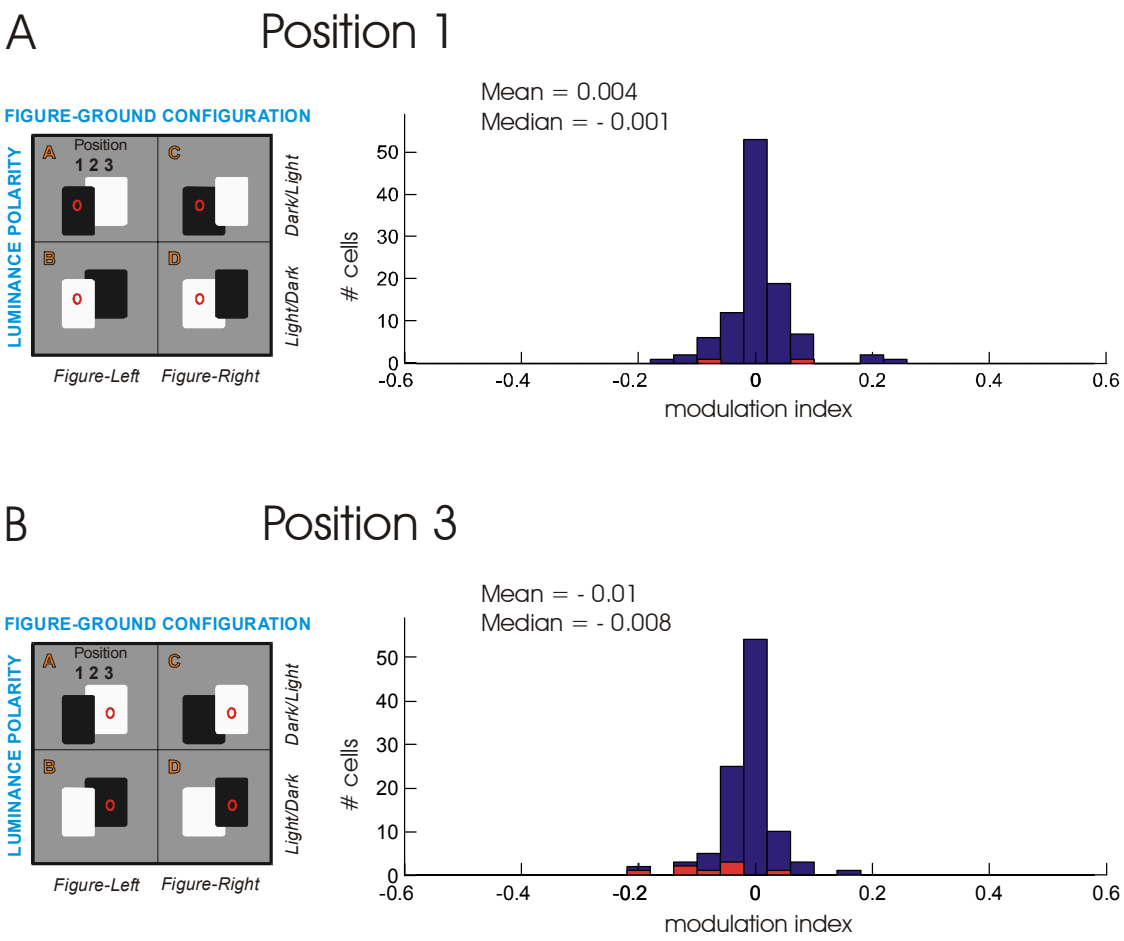

Figure 20: Modulation indices for all V1 cells tested with interiors of unambiguous figure-ground stimuli at two separate spatial positions. A: Indices for interior regions at position 1. In this position, if the index is > 0, the cell prefers figure over ground. B: Indices for interior regions at position 3. In this position, if the index is < 0, the cell prefers figure over ground.

Considering the small fraction of selective cells and their weak degree of selectivity, this difference may be due to chance or sampling error.

To study figure vs. ground effects in interior regions regardless of position, an additional set of modulation indices was calculated. This time the index was calculated as: $[(a\_f + b\_f + c\_f + d\_f) – (a\_g + b\_g + c\_g + d\_g)]/(a\_r + b\_r + c\_r + d\_r + a\_g + b\_g + c\_g + d\_g)$. a_f, b_f, c_f, and d_f represent the mean responses to all the figure rectangles depicted by the red receptive fields in A, B, C and D in Figure 21 on the left, and a_g, b_g, c_g, and d_g represent the mean responses to all the ground rectangles depicted by the green receptive fields. The distribution of indices is shown in the histogram on the right in Figure 21. 5% (5/103) of cells demonstrated a significant preference to an interior region regardless of position, with 80% (4/5) of the cells preferring figure over ground. The population (mean = 0.008, median = 0.004) is slightly shifted in the positive direction (t-test, $p = 0.014$), suggesting that when all the interior regions are grouped together regardless of position, some cells in V1 may have had a preference for figure over ground. However, overall, the results indicated that only a very small percentage (hovering around chance) of individual V1 cells demonstrated significant figure-ground selectivity in interior regions of unambiguous figure-ground stimuli, and the effects across the population were weak, as the majority of cells did not demonstrate any selectivity.

To compare, modulation indices at positions 1 and 3 in the ambiguous figure-ground stimuli are shown in Figures 22A and 22B, respectively. Both distributions are normal, and centered at zero (Jarque-Bera test of normality $p = 0.2803$, t-test, $p = 0.1728$, at position 1, Jarque-Bera test of normality, $p = 0.1815$, t-test, $p = 0.1445$, at position 3).

FIGURE 21

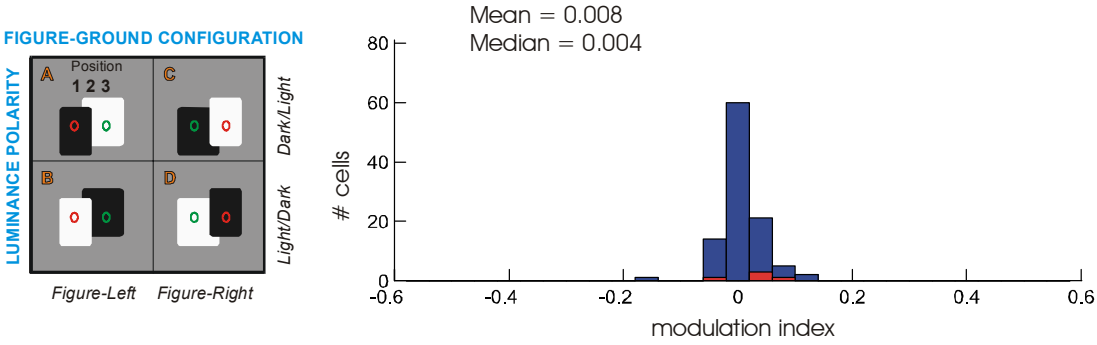

Figure 21: Modulation indices for all V1 cells tested with interiors of unambiguous figure-ground stimuli, regardless of spatial position. The modulation indices were calculated as described in the text.

FIGURE 22

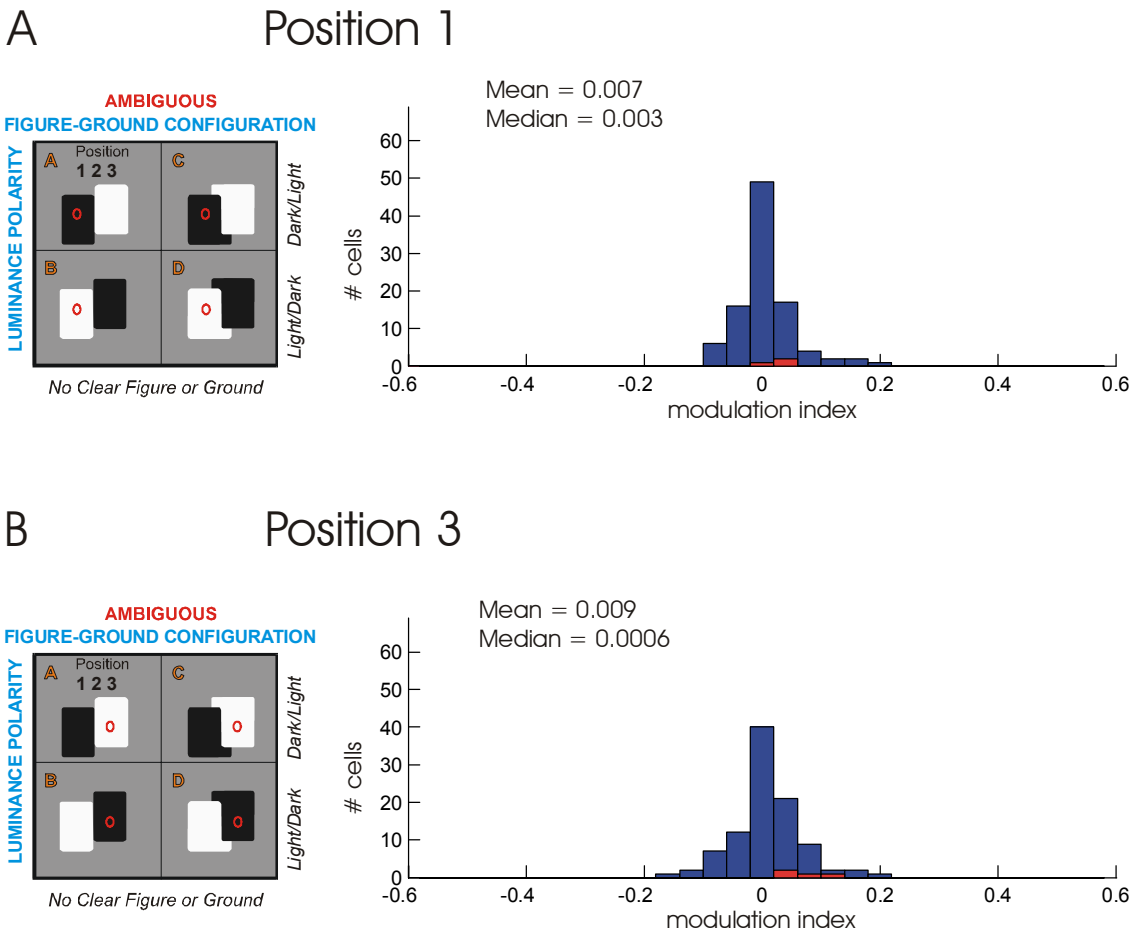

Figure 22: Modulation indices for all V1 cells tested with interiors of ambiguous figure-ground stimuli at two separate spatial positions. A: Indices for interior regions at position 1. 3% of cells demonstrated significant selectivity to interiors at this position. B: Indices for interior regions at position 3. 4% of cells demonstrated significant selectivity for interiors at this position. The number of individual cells demonstrating selectivity was smaller than that expected by chance (p= 0.05), and therefore were considered negligible. Both distributions were normally distributed and centered at zero, indicating no significant preference for the interior regions in the ambiguous figure-ground stimuli.

This confirms this population of cells did not demonstrate selectivity to figure-ground cues when their receptive fields were located in interior regions of the ambiguous figure-ground stimuli.

Overall, the results obtained in V1 suggested that a few cells had a significant preference to interior regions of unambiguous figure-ground stimuli, and that these cells preferred figure over ground more often than not. However, the percentage of cells with selectivity for the interiors was negligible, indicating these cells were not responding to the figure-ground cues or T-junction configuration located outside the classical receptive field.

V2 responses to interiors

We recorded from 170 cells in V2, and of those, all were tested with the interior regions of the unambiguous figure-ground stimuli, and 152 were additionally tested with the ambiguous figure-ground stimuli.

Figure 23-1 shows a V2 neuron's response to the interior regions of figure-ground rectangles. This cell demonstrated a significantly greater response ($p < 0.05$, 2-way ANOVA, Monte Carlo analyses) to the interior of figure regions over ground regions at position 1. Figure 23-1A depicts the cell's responses to dark interior regions. The cell's firing rate was significantly higher when its receptive field was located in the figure region as opposed to the ground region. Figure 23-1B shows the cell's responses to light interior regions. Again the cell's firing rate was significantly higher when its receptive field was located in the figure region. The averaged sum of responses in Figures 23-1A

FIGURE 23-1

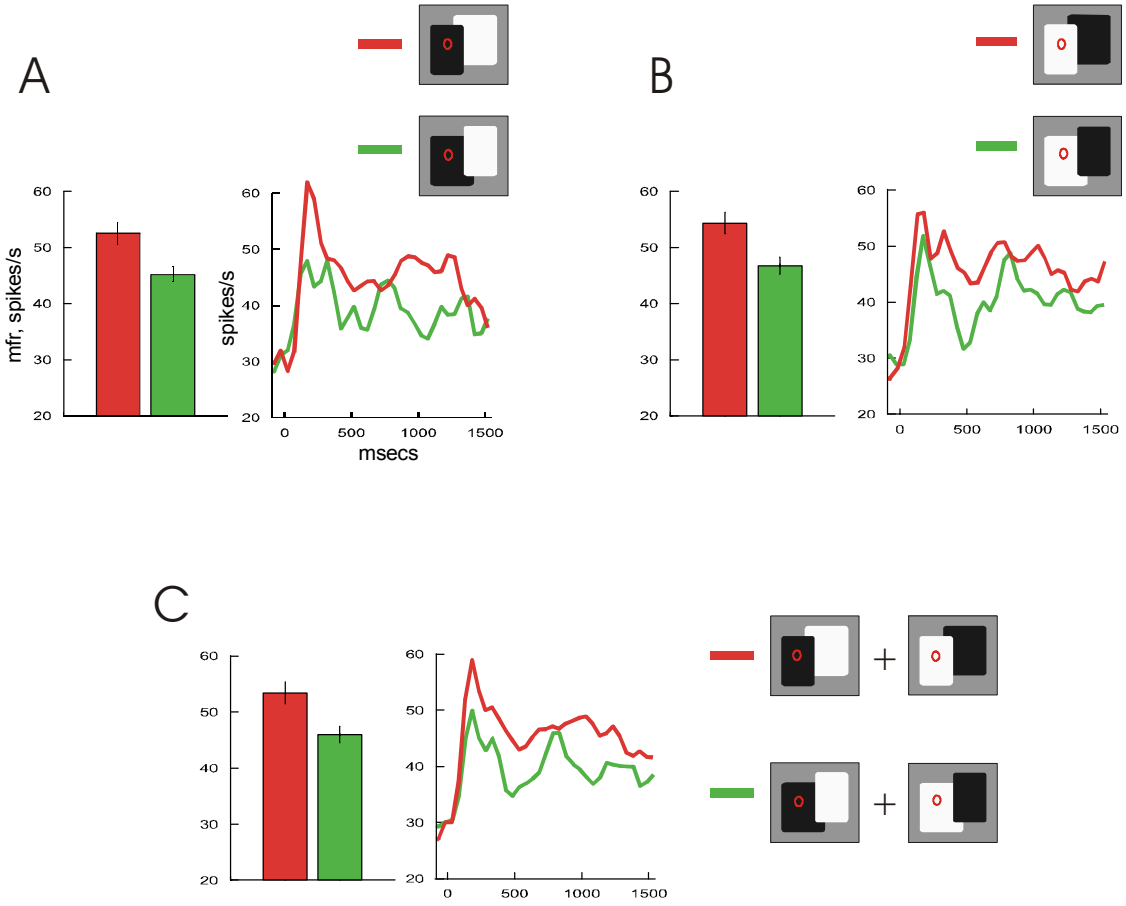

Figure 23-1: Responses from a single V2 cell demonstrating significant figure-ground selectivity in the interiors of unambiguous figure-ground stimuli. A: The responses of the cell to dark interior regions at position 1. This cell had a significantly greater response when its receptive field was located in the interior of a figure region as compared to a ground region. B: The responses of the cell to light interior regions as position 1. Again, this cell had a significantly greater response when its receptive field was located in the interior of the figure region. C: The averaged sum of the responses from A and B. This cell preferred conditions in which its receptive field was located in the figure regions, regardless of luminance.

and 23-1B are shown in Figures 23-1C. Overall, this cell preferred the interior of figure regions over ground regions at position 1.

Are there neurons which prefer figure over ground in interior regions regardless of position? The cell whose responses at position 1 were shown in Figure 23-1, appears to behave in this way. Its averaged responses to the interiors of figure regions vs. ground regions regardless of position are shown in Figure 23-2. Figure 23-2A presents the cell's averaged responses to the interiors of all figure regions and all ground regions at position 1. These graphs are the same as those depicted in Figure 23-1C. Figure 23-2B presents the cell's averaged responses to the interiors of all figure regions and ground regions at position 3. Again, this cell significantly preferred the interiors of figure regions over ground regions despite the luminance differences. Finally, Figure 23-2C presents the cell's averaged responses to all figure regions in both positions and its responses to all ground regions in both positions. This cell had a significantly greater response to the interiors of figure compared to ground, regardless of luminance and spatial location of the figure regions.

We additionally tested 152 of the 170 V2 cells with interior regions of ambiguous figure-ground stimuli. This revealed cells with selectivity for the interiors of one set of ambiguous figure-ground stimuli over the other. An example is shown in Figure 24. This cell demonstrated a significant preference ($p < 0.05$, 2-way ANOVA, Monte Carlo analyses) for the interior regions of ground rectangles compared to figure rectangles at position 1 (recall that the ambiguous figure-ground stimuli are comprised of either two figure rectangles or two ground rectangles). Figure 24A demonstrates the cell's

FIGURE 23-2

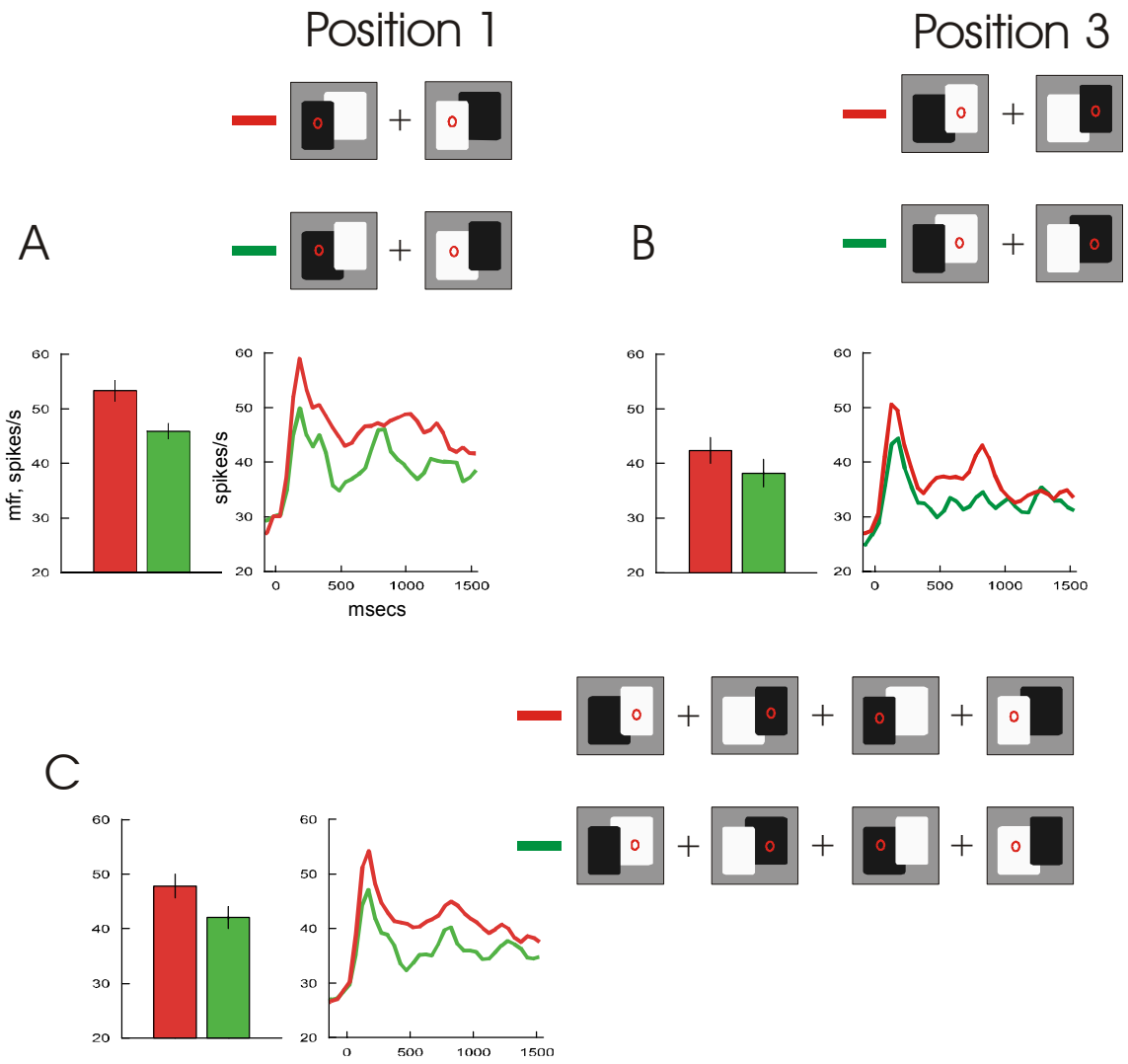

Figure 23-2: Responses from the same V2 cell in Figure 23-1, demonstrating significant figure-ground selectivity in the interiors of unambiguous figure-ground stimuli regardless of position. A: The averaged responses of the cell across dark and light regions to figure vs. ground at position 1. This cell had a significantly greater response when its receptive field was located in figure regions at this position regardless of luminance. B: The averaged responses of the cell across dark and light regions to figure vs. ground at position 3. Again, this cell had a significantly greater response when its receptive field was located in figure regions at this position regardless of luminance. C: The averaged sum of the responses from A and B. This cell preferred the conditions in which its receptive field was located in the figure regions, regardless of luminance or spatial position.

FIGURE 24

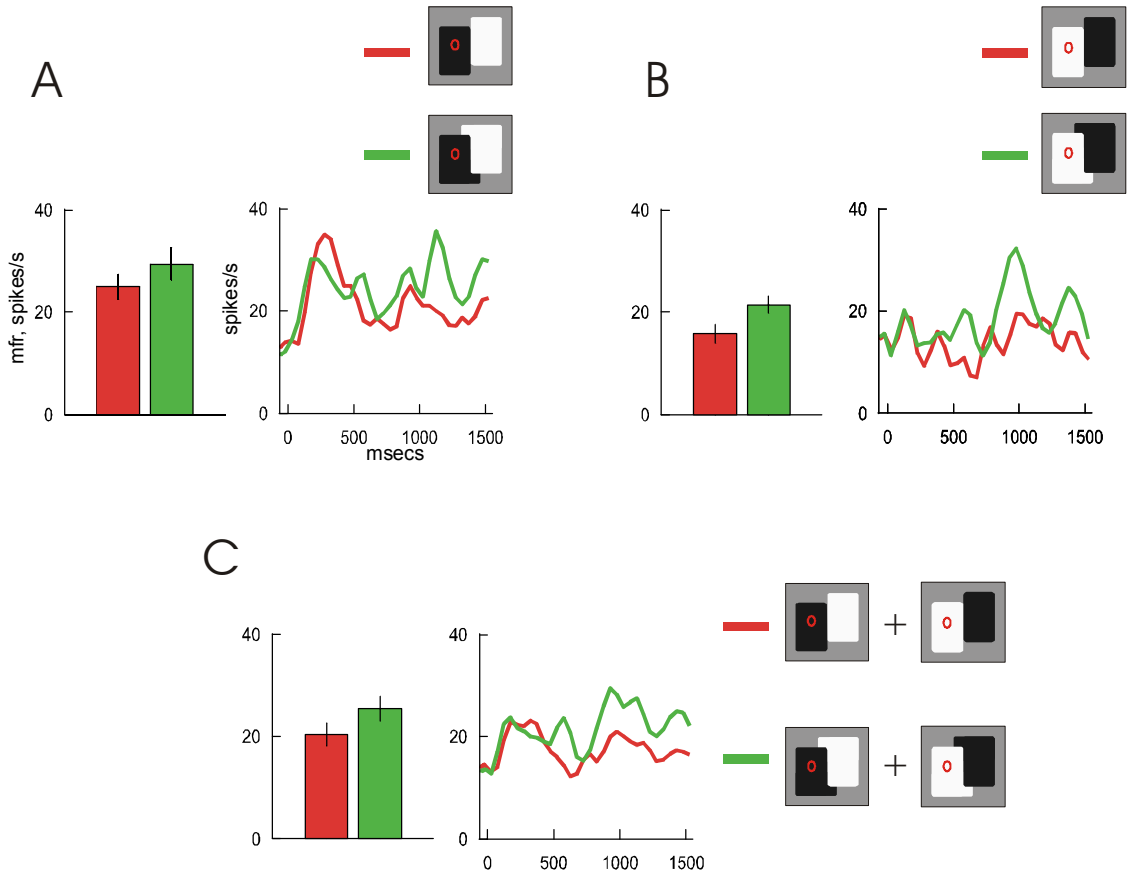

Figure 24: Responses from a V2 cell demonstrating significant figure-ground selectivity in the interiors of ambiguous figure-ground stimuli. A: The responses of the cell to dark interior regions at position 1. B: The responses of the cell to light interior regions at position 1. C: The averaged sum of the responses from A and B. This cell demonstrated a significant preference for the interiors of ambiguous figure-ground stimuli, indicating that the cell was sensitive to T-junction cues outside of the classical receptive field rather than figure vs. ground.

responses when its receptive field was located in the interior regions of dark rectangles. Its response was significantly greater when its receptive field was located in the interior of the ground rectangle as opposed to the figure rectangle. Similarly, Figure 24B shows that the cell significantly preferred the interior of the ground rectangle over the figure rectangle despite the change in luminance. Finally, the averaged sum of the responses from Figure 24A and 24B is shown in 22C. This plot indicates this cell had a significant preference for the interior region of ground rectangles over figure rectangles, regardless of the luminance of the rectangles. Figure-ground cues (T-junction configuration) outside the classical receptive field were apparently influencing this neuron's responses.

Over the neuronal population, not all cells which demonstrated significant selectivity for the interior regions did so at both positions in either the unambiguous or the ambiguous figure-ground stimuli. With unambiguous figure-ground stimuli, 10% (17/170 cells) demonstrated selectivity in interior position 1, and 8% (13/170) did so at position 3. 90% (27/30) of these selective cells preferred the interiors of figure rectangles to those of ground rectangles. 6% (10/170 cells) demonstrated significant selectivity regardless of position (i.e., at both positions 1 and 3), and similarly, of those, 90% preferred the interiors of figure rectangles over ground rectangles. As before, the predominant selectivity was for luminance. An average of 77% (125/170 at position 1, 138/170 at position 3) of cells preferred one luminance over the other. 5% (12/170 at position 1, 6/170 at position 3) demonstrated an interaction effect between luminance and figure-ground configuration.

With ambiguous figure-ground stimuli, 13% (19/152 cells) showed selectivity in the interior region at position 1, and 7% (10/152) showed selectivity in the interior region

at position 3. 9% (14/152) had selectivity for an interior region regardless of position. This is more than double the percentage found in V1. As in all other conditions, luminance selectivity was very strong amongst the cells, with an average of 78% of cells (115/152 at position 1, 123/152 at position 3) preferring one luminance over the other. 5% of cells (8/152 at position 1, 6/152 at position 3) demonstrated an interaction effect between the luminance and spatial configuration of the stimuli.

Modulation indices for the unambiguous figure-ground stimuli are shown in Figure 25. Figure 25A shows the distribution of indices at position 1, and Figure 25B shows the distribution of indices at position 3. At position 1, an index > 0 indicates the cell preferred the interior of figure rectangles. If the index is < 0, the cell preferred the interior of ground rectangles. The situation is reversed at position 3; if the index is < 0, the cell preferred the interior of figure, and if the index is > 0, the cell preferred the interior of ground. If the population of V2 cells showed a preference for figure over ground, the distribution of indices should be significantly shifted in the positive direction at position 1 and in the negative direction at position 3. Slight shifts did occur. In Figure 25A, at position 1, the population of indices (shown in blue) was not normally distributed (Jarque-Bera test of normality, $p = 0.0036$), and was weighted towards the positive direction with mean = 0.004 and median = 0.007 statistically greater than zero (signed rank test, $p = 0.0093$). For the cells which demonstrated significant selectivities by the ANOVAs and Monte Carlo analyses, the mean is also greater than zero (Jarque-Bera test of normality, $p = 0.5120$, t-test, $p = 1.5 \text{ X } 10e\text{-}4$). Of these cells, 94% (16/17) showed

FIGURE 25

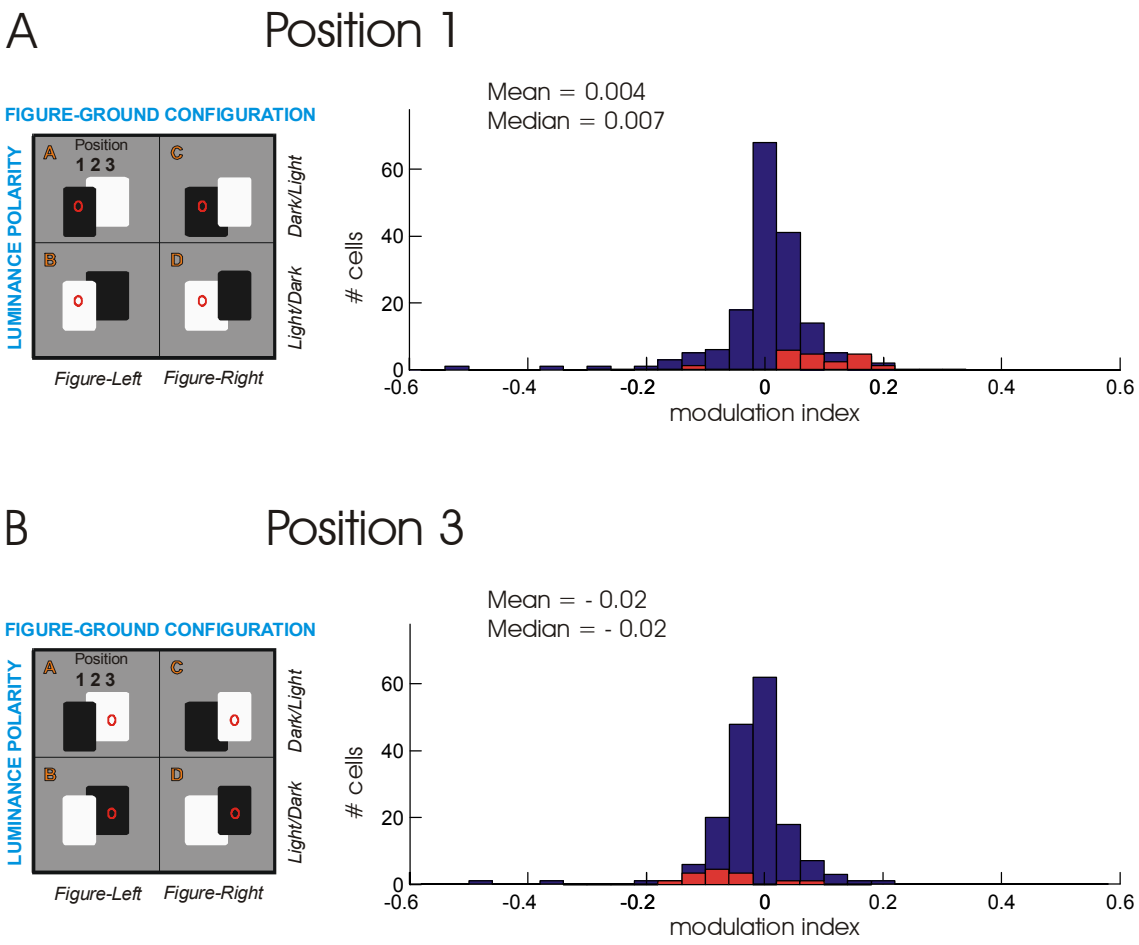

Figure 25: Modulation indices for all V2 cells tested with interiors of unambiguous figure-ground stimuli at two separate spatial positions. The indices for all cells in the population are shown in blue, and the indices for selective cells are shown in red. A: Indices for interior regions at position 1. In this position, if the index is > 0, the cell prefers figure over ground. 10% of cells demonstrated significant figure vs. ground selectivity at this position. As a population, the distribution was not normally distributed, and shifted significantly in the positive direction, indicating overall, these cells preferred figure over ground. B: Indices for interior regions at position 3. In this position, if the index is < 0, the cell prefers figure over ground. 8% of cells demonstrated significant figure vs. ground selectivity at this position. Again, the distribution was not normally distributed, and its mean was shifted significantly in the negative direction, indicating the population preferred figure over ground. Approximately twice the percentage of individual V2 cells demonstrated a preference to interior regions as compared to V1 cells. Of those, 90% preferred figure over ground.

preference for the interior of figure regions over those of ground regions. The same effect was present at position 3, shown in Figure 25B. Again, the population of indices was not normally distributed (Jarque-Bera test of normality, $p = 0.00031$), and skewed towards the negative direction, with mean = -0.02, and median = -0.02 statistically less than zero (signed rank test, $p = 5.5 \text{ X } 10e\text{-}7$). The grouped indices of the individual cells that demonstrated figure-ground selectivity, had a mean less than zero as well (Jarque-Bera test of normality, $p = 0.6026$, t-test, $p = 0.003$), and 85% (11/13) of them preferred the interiors of figure regions over those of ground regions. However, even though there are hints that the population of neurons may prefer figure over ground, the effect is weak since the means and medians of these distributions are close to zero. Rather, it is possible that these distributions may be biased by individual neurons which demonstrated a significant preference for figure over ground.

Modulation indices grouping all figure regions and ground regions are shown in Figure 26. The overall distribution was normal, but shifted towards the positive direction, (Jarque-Bera test of normality, $p = .5383$, t-test, $p = 7.3 \text{ X } 10e\text{-}5$) indicating that this group of cells may have a preference for figure over ground in interior regions regardless of position. Of the individual cells which demonstrated significant selectivity to the interior regions regardless of position, 90% (9/10) preferred figure over ground. However, the skew of the distribution in the positive direction was weak, as the mean and median were very close to zero, indicating that the majority of cells did not demonstrate figure-ground selectivity.

Modulation indices for the population of cells tested with ambiguous figure-ground stimuli are shown in Figure 27. Figure 27A shows the indices at position 1, and

FIGURE 26

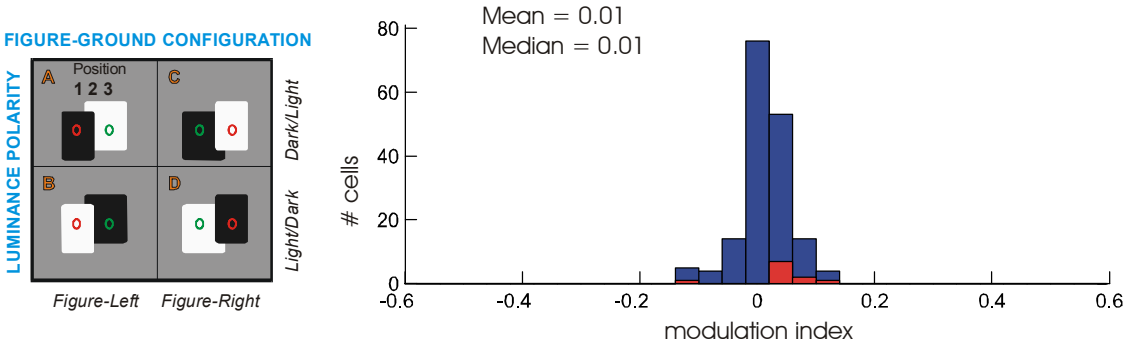

Figure 26: Modulation indices for all V2 cells tested with interiors of unambiguous figure-ground stimuli, regardless of spatial position. If the index is > 0, the cell prefers figure over ground. 6% of cells demonstrated significant selectivity for the interior regions regardless of position, and again, 90% of them preferred figure over ground. The indices were normally distributed, but its mean was significantly greater than zero, suggesting that the population demonstrated a preference for figure over ground, regardless of spatial position.

FIGURE 27

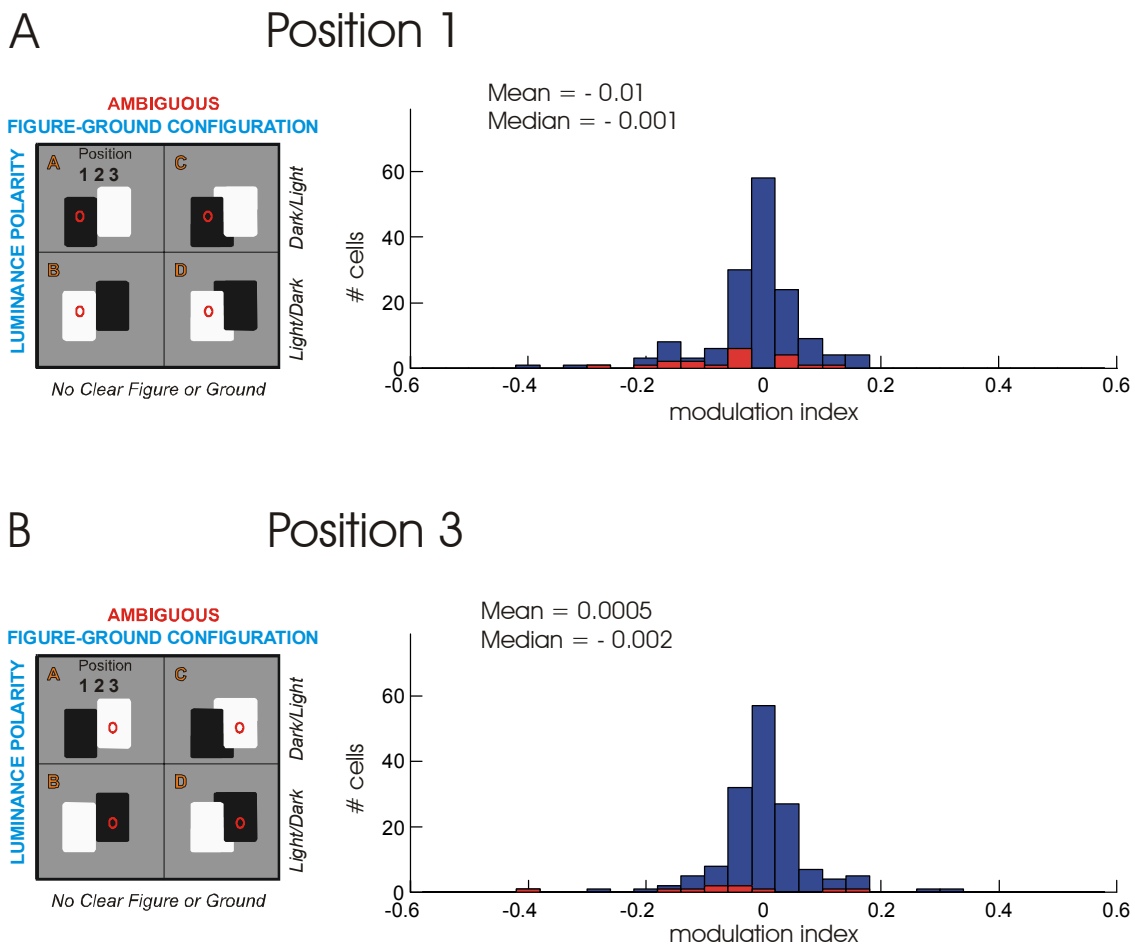

Figure 27: Modulation indices for all V2 cells tested with interiors of ambiguous figure-ground stimuli at two separate spatial positions. At either position, if the index is > 0, the cell prefers the interiors of the figure rectangles, and if the index is < 0, the cell prefers the interiors of the ground rectangles. A: Indices for interior regions at position 1. 13% of cells demonstrated significant selectivity for interiors regions at this position. B: Indices for interior regions at position 3. 7% of cells demonstrated significant selectivity for interior regions at this position. At both positions, the distributions of indices were not normally distributed, but they were centered about zero. This indicates these cells did not have a significant preference for either the figure rectangles or the ground rectangles in the ambiguous figure-ground stimuli.

Figure 27B shows the indices at position 3. If a cell's index is > 0 at either position, the cell preferred the interiors of figure rectangles over those of ground rectangles in the ambiguous figure-ground stimuli. Similarly, if a cell's index is < 0 at either position, the cell preferred the interior of ground rectangles over those of figure rectangles. At both positions 1 and 3, the distributions were not normally distributed (position 1, Jarque-Bera test of normality, $p = 0.001$, position 3, Jarque-Bera test of normality, $p = 2.5 \text{ X } 10e\text{-}6$). However, their medians were also not significantly different than zero (position 1, signed rank test, $p = 0.2846$, position 3, singed rank test, $p = 0.8758$), indicating the population of cells was not selective for interiors of rectangles at either position. These distributions were different than those obtained for the unambiguous figure-ground stimuli, in which the populations were slightly skewed in preference of figure regions over ground regions. This result lends additional support to the hypothesis that cells do exist which prefer figure over ground, and that separate populations of cells may be responding to figure vs. ground and to figure-ground cues.

When tested with unambiguous figure-ground stimuli, the fraction of V2 cells that responded with significant preference to interior regions was about twice that found in V1. Perhaps more interestingly, the vast majority of these selective cells preferred the interiors of figure regions over ground regions. The population mean was slightly biased towards a figure-ground preference, perhaps due to the influence of the individually selective cells. Results from the V2 cells tested with ambiguous figure-ground stimuli indicated that there also existed individual V2 cells which had significant selectivity for these interior regions. The responses of these cells may have been influenced by figure-ground cues located outside the classical receptive field. As a population, these cells did

not appear to be slightly biased towards one set of ambiguous figure-ground stimuli over the other, differing somewhat from the results obtained from the unambiguous stimuli. Furthermore, separate groups of cells demonstrated preferences for the unambiguous figure-ground stimuli and the ambiguous figure-ground stimuli. Because groups of cells appeared to behave differently when tested with the two different sets of stimuli, this lends further support to the hypothesis that separate cells encode figure vs. ground and figure-ground cues.

CHAPTER 4

Discussion

Our experiments were designed to examine whether neurons in V1 or V2 demonstrated border ownership selectivity or figure vs. ground selectivity in interiors—or whether neurons were merely responding to figure-ground cues (T-junction configuration) located outside of the classical receptive field. We began our studies in an awake, fixating monkey. We found little evidence in V1 to suggest it played a significant role in figure-ground interpretation. However, we found cells in V2 which were selective for border ownership at edges and figure vs. ground in interior regions when the cells were tested with unambiguous figure-ground stimuli. We also found cells which were selective for interior regions when they were tested with ambiguous figure-ground stimuli. These results suggested that V2 may play a stronger role than V1 in the determination of figure-ground, and additionally that cells may exist in V2 which respond to figure-ground cues.

We pursued further experiments in which we trained two additional monkeys on a behavioral task which ensured their attention to the spatial configuration of the stimuli. In V1, some individual cells demonstrated border ownership selectivity in unambiguous figure-ground stimuli, and others demonstrated selectivity to figure-ground cues in ambiguous figure-ground stimuli. However, these cells were few and their selectivities were weak. The population as a whole did not demonstrate appreciable selectivity at the borders in either the unambiguous or the ambiguous figure-ground stimuli. In V2, individual cells demonstrating significant selectivity for border ownership or figure-

ground cues were encountered twice as frequently as in V1, although the population of V2 cells also did not appear to have appreciably selectivity for borders in either set of stimuli.

In interior regions in V1, we found negligible selectivity amongst individual cells and across the population for figure vs. ground in unambiguous stimuli or for figure ground cues in ambiguous stimuli. The population of cells also did not demonstrate selectivity in the interiors in either set of stimuli. However, in V2, we did find cells with selectivity for interior regions in the unambiguous figure-ground stimuli, and the vast majority of these cells preferred figure over ground. The distributions of modulation indices were biased slightly towards the preference of figure regions, perhaps due to the presence of these figure-selective cells. Because these populations were centered very close to zero, the majority of cells did not demonstrate appreciable selectivity, though it is clear that a small group of cells in V2 did. Additionally, we found cells in V2 with selectivity for the interior regions of the ambiguous figure-ground stimuli, again consistent with the idea that certain cells respond to figure-ground cues. The modulation indices indicated the population as a whole was not selective for these cues, though.

Although the majority of cells tested in V1 and V2 did not demonstrate appreciable border ownership or figure-ground selectivity, we did find individual cells in these areas which did. The effects were stronger in V2, with approximately twice the percentage of cells showing selectivity. Here, the cells which demonstrated a preference for border ownership or figure vs. ground in the unambiguous figure-ground stimuli were, by and large, separate from those which demonstrated a preference for figure-ground cues in the ambiguous figure-ground stimuli. This suggests that separate

populations of cells may respond to the true figure-ground configuration and to figure-ground cues. Because the distributions of modulation indices tested with unambiguous stimuli were statistically different from those obtained with ambiguous stimuli, this further suggests that certain cells respond to the true figure-ground configuration of stimuli, whereas others respond to figure-ground cues.

An alternate interpretation of our results would suggest that cells may be merely responding to particular stimulus configurations. For instance, certain cells may prefer a particular arrangement of light dark rectangles, regardless of any figure-ground aspects or cues in the stimuli. While this interpretation cannot be ruled out, we felt it reasonable to analyze our results using figure-ground configuration and figure-ground cues as the determining factors in our stimuli.

Comparison to other studies

Our experiments build upon and extend the results of previous studies which examined border ownership selectivity and contextual modulation that may give rise to figure-ground representation. Our experiments differ from previous studies in two ways. First, we asked animals to perform a behavioral task that ensured their attention to the stimuli. Second, we studied the responses of V1 and V2 neurons to borders and interior regions in both unambiguous and ambiguous (control) figure-ground stimuli. In general, our findings seem consistent with previous results on border ownership selectivity in V1 and V2. Our results are also consistent with the most recent previous study of figure vs. ground selectivity in V1. Previous studies did not examine figure vs. ground selectivity in V2.

Baumann et al. (1997) studied border ownership selectivity in areas V1 and V2 using stimuli in which figure-ground occlusion cues, in this case, terminating grating lines, were located within the classical receptive field. They found no neurons in V1 which demonstrated border ownership selectivity, while 22% of neurons in V2 showed this property. We found 10% in V1 and 21% in V2, but it is difficult to compare the percentages in these two experiments because substantially different stimuli were used. Furthermore, our experiments used figure-ground cues located outside the classical receptive field, whereas their experiments used figure-ground cues within the classical receptive field. Finally, their criteria for selectivity were based on a cutoff value for firing indices, whereas ours were based on the outcomes of ANOVAs. However, both our studies found cells in V2 which demonstrated border ownership selectivity. Our additional finding of cells in V1 with this property do not necessarily contradict the report of Baumann et al., in light of the differences in experimental design, as well as potential sampling differences.

Another study, more similar in design to ours, was performed by Zhou et al. (2000). They also used contrast-defined figures as their stimuli. In their primary experiments, they recorded V1 and V2 cells' responses at the borders of uniformly colored squares on uniformly colored backgrounds. Using ANOVAs, they found 3% of cells in V1 and 15% of cells in V2 which demonstrated border ownership selectivity regardless of luminance polarity. They additionally recorded from a few cells (8 total in V1, 16 in V2) using overlapping rectangles, similar to the ones we used in our experiment. They did not examine the effect of changing the luminance polarity, though, so it is unfortunately impossible to directly compare results. However, our results are

consistent with those of Zhou et al., as we also found only a few neurons in V1 which were selective for border ownership, and a larger number in V2. Overall, the percentages of neurons demonstrating a preference for border ownership in all of our studies are very similar, in spite of the differences in experimental design.

Previous studies on figure-ground selectivity in interior regions have yielded inconsistent results. In 1995, Lamme reported that nearly all the V1 neurons he sampled preferred the interior of orientation-defined figure regions as compared to similarly oriented but uniform background regions. His figures consisted of a square of hashed lines against a backdrop of perpendicularly oriented hashed lines; the uniform background regions, with which he compared the figural responses, were oriented hashed lines. Zipser et al. (1996) built on Lamme's experiments by varying the size of the figure to determine if the V1 neurons' enhanced responses would persist across various spatial scales. He claimed that they did, in approximately 40% of the neurons he tested. Lee et al. (1998) reported results similar to those of Zipser, and additionally found the effect in contrast-defined stimuli, i.e., when uniformly colored figures were used. However, Rossi et al. (2001), using orientation-defined figures, refuted Zipser et al's and Lee et al's report of figure selectivity over a large range of spatial scales. Rossi et al. determined that their V1 neurons gave enhanced responses only when the figure's border was very close to the edge of the cells' receptive fields. This suggested that only the cues in the local environment around the receptive field, or texture boundaries, influenced the neurons' firing rate, rather than true figure-ground configuration.

In V1 we found negligible selectivity for interior regions in both unambiguous as well as ambiguous stimuli. However, we did not systematically study the effect of spatial

scale, and our stimuli were substantially larger than the receptive fields of the V1 neurons. Our results, therefore, may be consistent with those of Rossi et al., which found no enhanced neuronal responses in V1 when the figure-ground texture boundaries were located at a distance outside of the classical receptive field. Our results in V2, however, indicate the existence of figure-ground selective cells there. No other previous study had looked at neuronal responses to the interiors of figure regions in V2.

Our experiments go farther than the previous ones in examining border ownership selectivity at boundaries as well as figure-ground selectivity in interiors by using ambiguous (control) as well as unambiguous figure-ground stimuli. The ambiguous figure-ground stimuli contained the same figure-ground cues, but no genuine figure-ground configuration. These stimuli revealed that separate cell populations in V1 and V2 respond to border ownership and to figure-ground cues in the ambiguous stimuli. Therefore, different cells apparently encode these properties. Similar arguments apply for responses to interior regions in area V2.

Furthermore, because our animals had to perform a behavioral task which required them to visually discriminate between two very similar stimuli, and hence pay close attention to their spatial configuration, our experiments also examined whether there were attentional effects in the cells' responses. Apparently such effects did not exist, as the fraction of cells selective for border ownership or figure vs. ground in interiors was not enhanced. The percentages of neurons which demonstrated preferences were similar in the fixating monkey and in the match-to-sample monkeys. This may indicate that attention may not be required in order for these selectivities to occur.

Overall, our results support the previous findings of border ownership in areas of visual cortex, although we found slightly more cells with this property in V1, but approximately the same amount in V2 percentage-wise. While we did not find figure-ground selectivity in interior regions in V1, this could be due to the relatively larger stimuli we used. We did, however, find figure-ground selectivity in interior regions in V2, which had not been found previously. We additionally found that figure-ground cues, when located outside the classical receptive field, can also enhance neurons' responses regardless of the actual figure-ground percept of the stimuli, suggesting that separate neurons may respond to these cues rather than figure vs. ground per se. Finally, we found that attention does not appear to enhance neurons' responses to figure vs. ground, suggesting that figure-ground discrimination may be a lower-order process.

Conclusions and Directions for Future Work

Our experiments focussed exclusively on the role of occlusion cues in figure-ground discrimination. A small percentage of cells in V1 showed border ownership selectivity and a few others responded selectively to figure-ground cues at borders. There was no evidence that cells in V1 were selective for figure or ground or figure-ground cues when a cell's receptive field was located within an interior region. In V2, a larger fraction of cells showed border ownership selectivity, and some had significant selectivity for the figure region of figure-ground stimuli. A separate population of V2 cells apparently showed selectivity for figure-ground cues at edges and in interior regions.

Further electrophysiology experiments could help reveal how the visual system determines figure and ground. For instance, it would be useful to know what the receptive field structures of the selective cells are in V1 and V2. In Hubel and Wiesel's (1968) terms, could the cells be classified as simple, complex, or hypercomplex? Knowing the answer to this question would be a useful first step to approaching the question of what wiring schemes might explain figure-ground behavior. One might speculate, for example, that cells with selectivity for the presence of t-junctions might receive excitatory input from hypercomplex cells that respond to the presence of corners in their receptive fields. The cells that respond selectively to the interior region of the figure in figure-ground stimuli might in turn receive input from cells that respond selectively to the presence of t-junctions. Further questions arise: is there an anatomical arrangement of figure-ground selective cells in V1 and V2? Are there "columns" or "blobs" of these cells, or are they simply interspersed among other types of neurons?

Additional experiments could utilize a variety of other cues, such as stereo or color, to determine whether figure-ground selective cells are selective regardless of the nature of the figure-ground cues, as Zipser et al. suggested in V1. If, instead, it turns out that separate populations of cells encode for different kinds of figure-ground cues in V1 and V2, how and where in the brain is the information about these different cues integrated? The integration might involve convergent connections from cells in V1 and V2 with the different selectivities onto neurons in some higher center. In locating this center and defining the wiring it would be useful to know where these selective cells project. Because it is possible that the figure-ground task involves many different cells, it may be worthwhile to utilize other techniques which involve sampling greater numbers of

cells. For approaching these problems, anatomical tracer studies, fMRI imaging, and multi-electrode recordings would be useful complements to single-electrode recordings.

References:

Allman, J., F. Miezin, and E. McGuinness (1985). “Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT).” Perception. 14: 105-26.

Baumann, R., R. van der Zwan, and E. Peterhans (1997). "Figure-ground segregation at contours: a neural mechanism in the visual cortex of the alert monkey." European Journal of Neuroscience. 9: 1290-303.

Dobkins, K. R. and T. D. Albright (1994). "What happens if it changes color when it moves?: the nature of chromatic input to macaque visual area MT." Journal of Neuroscience. 14: 4854-70.

Gattass, R., C. G. Gross, and J. H. Sandell (1981). "Visual topography of V2 in the macaque." Journal of Comparative Neurology. 201: 519-39.

Gilbert, C. D. and T. N. Wiesel (1990). "The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat." Vision Research. 30: 1689-701.

Heitger, F., R. Von der Heydt, E. Peterhans, L. Rosenthaler, O. Kubler (1998). "Simulation of neural contour mechanisms: representing anomalous contours." Image and Vision Computing. 16: 409-423.

Hubel, D. H. and T. N. Wiesel (1968). "Receptive fields and functional architecture of monkey striate cortex." Journal of Physiology. 195: 215-43.

Hubel, D. H. and T. N. Wiesel (1977). "Ferrier lecture. Functional architecture of macaque monkey visual cortex." Proceedings of the Royal Society of London - Series B: Biological Sciences. 198: 1-59.

Judge, S. J., B. J. Richmond, and F. C. Chu (1980). "Implantation of magnetic search coils for measurement of eye position: an improved method." Vision Research. 20: 535-8.

Kastner, S., P. De Weerd, and L. G. Ungerleider (2000). "Texture segregation in the human visual cortex: A functional MRI study." Journal of Neurophysiology. 83: 2453-7.

Kastner, S., H. C. Nothdurft, and I. N. Pigarev (1997). "Neuronal correlates of pop-out in cat striate cortex." Vision Research. 37: 371-6.

Knierim, J. J. and D. C. van Essen (1992). "Neuronal responses to static texture patterns in area V1 of the alert macaque monkey." Journal of Neurophysiology. 67: 961-80.

Koffka, K. (1935). Principles of gestalt psychology. New York, Harcourt, Brace, and World.

Lamme, V. A. (1995). "The neurophysiology of figure-ground segregation in primary visual cortex." Journal of Neuroscience. 15: 1605-15.

Lee, T. S., D. Mumford, R. Romero, and V. A. Lamme (1998). "The role of the primary visual cortex in higher level vision." Vision Research. 38: 2429-54.

Li, W., P. Thier, and C. Wehrhahn (2000). "Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys." Journal of Neurophysiology. 83: 941-54.

Li, W., P. Thier, and C. Wehrhahn (2001). "Neuronal responses from beyond the classic receptive field in V1 of alert monkeys." Experimental Brain Research. 139: 359-71.

MacEvoy, S. P., W. Kim, and M. A. Paradiso (1998). "Integration of surface information in primary visual cortex." Nature Neuroscience. 1: 616-20.

Maffei, L. and A. Fiorentini (1976). "The unresponsive regions of visual cortical receptive fields." Vision Research. 16: 1131-9.

Nelson, J. I. and B. J. Frost (1978). "Orientation-selective inhibition from beyond the classic visual receptive field." Brain Research. 139: 359-65.

Peterhans, E. and F. Heitger (2001). "Simulation of neuronal responses defining depth order and contrast polarity at illusory contours in monkey area V2." Journal of Computational Neuroscience. 10: 195-211.

Rossi, A. F., R. Desimone, and L. G. Ungerleider (2001). "Contextual modulation in primary visual cortex of macaques." Journal of Neuroscience. 21: 1698-709.

Rossi, A. F., C. D. Rittenhouse, and M. A. Paradiso (1996). "The representation of brightness in primary visual cortex.[comment]." Science. 273: 1104-7.

Von der Heydt, R. and E. Peterhans (1989). "Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity." Journal of Neuroscience. 9: 1731-48.

Von der Heydt, R., E. Peterhans, and G. Baumgartner (1984). "Illusory contours and cortical neuron responses." Science. 224: 1260-2.

Worgotter, F., W. J. Daunicht, and R. Eckmiller (1986). "An on-line spike form discriminator for extracellular recordings based on an analog correlation technique." Journal of Neuroscience Methods. 17: 141-51.

Zhou, H., H. S. Friedman, and R. Von der Heydt (2000). "Coding of border ownership in monkey visual cortex." Journal of Neuroscience. 20: 6594-611.

Zipser, K., V. A. Lamme, and P. H. Schiller (1996). "Contextual modulation in primary visual cortex." Journal of Neuroscience. 16: 7376-89.

PART II

Structure-from-Motion in MT

CHAPTER 5

Encoding of three-dimensional structure-from-motion by primate area MT neurons

David C. Bradley, Grace C. Chang and Richard A. Andersen

Abstract

We see the world as three-dimensional. However, because the retinal image is flat, we must derive the third dimension—depth—from two-dimensional cues. Image movement provides one of the most potent cues for depth (1-6). For example, the shadow of a contorted wire appears flat when the wire is stationary, but rotating the wire causes motion in the shadow, which suddenly appears three-dimensional. The neural mechanism of this effect, known as "structure-from-motion," has not been discovered. We studied area MT, a primate cortical region known to be involved in visual motion perception. Two rhesus monkeys were trained to fixate their gaze while viewing two-dimensional projections of transparent, revolving cylinders. These stimuli are like Necker cubes in that they appear three-dimensional, but the surface order one perceives (front vs. back) tends to reverse spontaneously. These reversals occur because the stimulus does not specify which surface is in front or back. Monkeys reported which surface order they perceived after viewing the stimulus. In many of the neurons tested, there was a reproducible change in activity coinciding with reversals of the perceived surface order, even though the stimulus remained identical. This suggests that MT has a basic role in structure-from-motion perception.

We trained two rhesus monkeys to fixate a stationary target while we showed the two dimensional projection of a revolving, random-dot cylinder (Figure 28A). This projection contains opposite-going motions which convey a sense of front and back, or *surface order*. Monkeys then reported the direction of the front surface by glancing at one of two targets that appeared on either side of the cylinder's former position. Since two dimensional projections are flat, they do not specify the surface order, so the monkeys' answers reflect their three dimensional perception of the stimulus.

We also showed rotating cylinders whose structure (and thus surface order) was specified with disparity. Some of these were flattened by multiplying the disparity (depth) of each dot by a fraction (0, 12.5, 25, 50 or 100%). All cylinders had their center at zero disparity, so one surface appeared near, the other far, relative to the fixation depth. Figure 28B shows that performance—the ability to judge surface order—decreased predictably as the disparity decreased, suggesting that monkeys were doing the task as required.

In simultaneous, single-neuron recordings of MT activity, we oriented cylinders such that one surface moved in the neuron's preferred direction (determined in preliminary tests), the other in the opposite direction. We expected responses to depend on surface order because MT cells tend to prefer motion either behind or in front of the fixation point (far or near) (7). Preferred-direction motion on the "active" side tends to excite, while antipreferred motion on the active side tends to suppress (8). Therefore, one of the two surface orders should be optimal because it places preferred motion on the active side while placing antipreferred motion on the other side. Indeed, when the highest-disparity cylinders were shown, 68/109 MT cells responded significantly better to

FIGURE 28

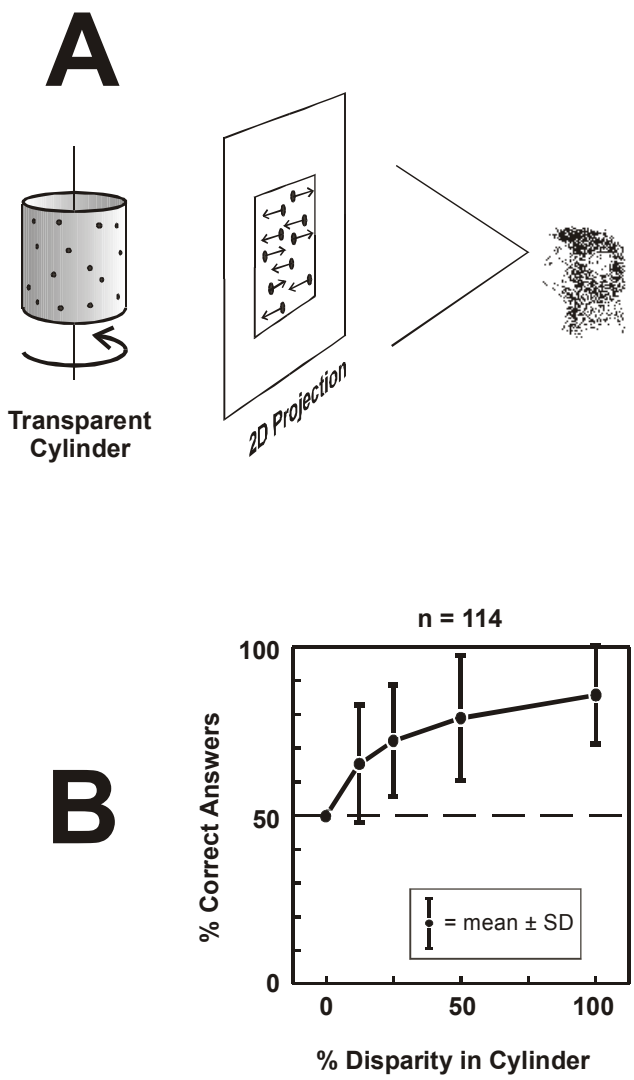

Figure 28: Panel A: Monkeys viewed two-dimensional projections of three-dimensional, revolving cylinders, then reported the direction of rotation they perceived. Panel B: Psychometric function (all data pooled and averaged) relating performance to the amount of disparity in the cylinders. The abscissa shows disparity as a percentage of the disparity one would see looking at a real (3D) cylinder. The ordinate shows the percentage of correct answers regarding the surface order. Error bars show standard deviation.

one of the 2 surface orders ($p<0.05$, t-test).

Similar response differences were linked to the *perception* of surface order. The cell in Figure 29, for example, preferred the front going down, the back up, and this was true whether that surface order was specified with disparity (top) or simply perceived as such in a zero-disparity stimulus (bottom). Of the 68 cells that responded preferentially to a given, disparity-specified surface order (see above), 34 responded differently when a *given* stimulus was *perceived* with different surface orders ($p<0.05$, t-test; see Methods). Most cells (27/34) showed "correlated" behavior, meaning responses were higher when the neuron's preferred order (defined at the highest disparity) was perceived, and this was true for cells that responded maximally when their preferred direction was in front (17/20) as well as those favoring their preferred direction in back (7/9; 5 cells not classifiable as near or far). The importance of this is discussed below. Given the low frequency of cells with the opposite, "anticorrelated" behavior (7 of 68 possible), it is not clear whether a distinct cell class of this type really exists.

Although disparity cues bias perception in favor a particular surface order, all stimuli were potentially bistable (see Figure 28B). Figure 30A shows that whatever the disparity, and whatever the specified surface order, responses were higher when the neuron's preferred surface order was perceived (vs. the non-preferred order). Moreover, whether the variable in question was the specified surface order or the perceived surface order, the time course over which activity diverged was similar (Figure 30B). Thus, to the extent that the perceived surface order of a given stimulus differed, MT activities tended to reflect that difference.

FIGURE 29

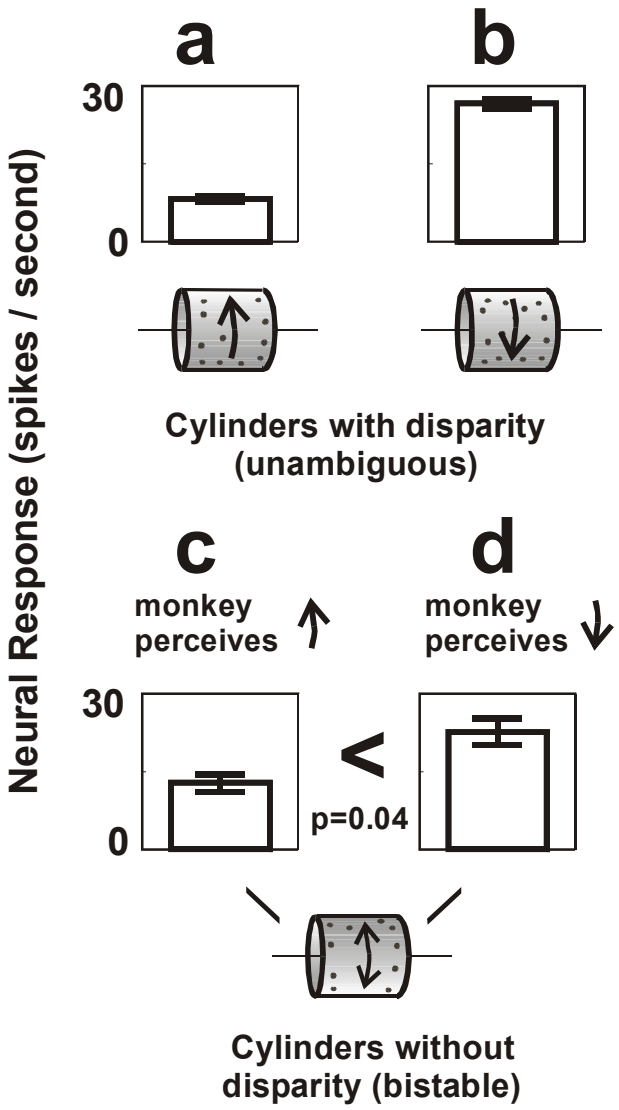

Figure 29: Data from an MT neuron. Panels A and B: responses to 100%-disparity cylinders (using only trials with correct responses). Panels C and D: responses of the same cell to zero-disparity cylinders, perceived as going front-up (panel C) and front-down (panel D).

FIGURE 30

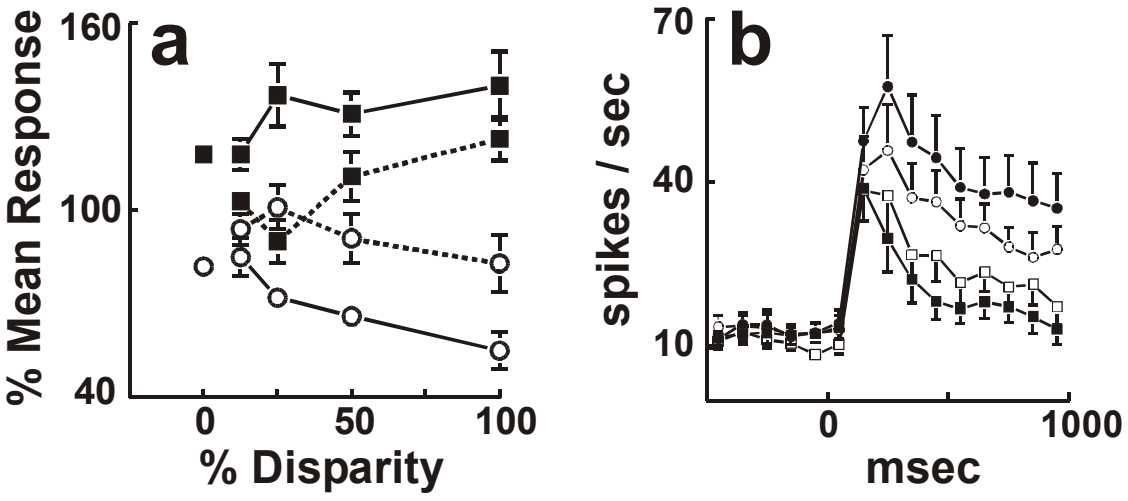

Figure 30: Averaged data from "correlated" MT cells (n=27). Panel A: Perceptual effects at different disparities. Responses were normalized to the mean response for each disparity and plotted as a function of disparity. Data show mean responses ± standard error. Upper 2 curves (squares): Actual surface order matches neuron's preferred surface order. Lower 2 curves (circles): actual surface order is opposite to neuron's preferred order. Solid lines: correct trials. Broken lines: error trials. At zero disparity, since there is no actual surface order (the stimulus has zero disparity), the data points represent the two perceived surface orders. Panel B: Time course of responses at zero disparity (inner traces) and 100% disparity (outer traces; correct trials only). Circles represent the neurons' preferred surface order (perceived or actual); squares represent the non-preferred order. The stimulus was visible from time 0 to time 1000 msec. Data show mean response ± standard error (in some cases error bars are smaller than the symbols).

Results

Previous experiments with flat patterns showed that opposite motion directions suppress MT responses, but this suppression decreases, and in some cases changes to facilitation, when opposite directions are shown at different disparities (11, 8). This suggests inhibitory connections between MT cells tuned for opposite directions and similar depths (Figure 31A), and excitatory connections between cells tuned for opposite directions and different depths (i.e., near vs. far). Such depth-dependent interactions may be important for computing surface movement because they emphasize coherent (same-direction) motion signals while suppressing random signals (motion noise) from a given surface (8, 12).

Structure-from-motion perception may begin with the bistable nature of this circuitry. MT cells typically prefer either near or far stimuli, but their tuning is broad enough that they also respond to zero-disparity stimuli (7). Therefore, a zero-disparity cylinder projection could potentially activate 4 neuronal pools, tuned (assuming a vertical cylinder) for near-right, near-left, far-right, and far-left (Figure 31A). But because of the inhibition and excitation discussed above, an even distribution of activity would be unstable, tending to "fall" into a distribution that places opposite directions in different depth channels. For example, an increase in the activity of near-right cells could lead to a suppression of near-left cells and an activation of far-left cells; the far-left cells, in turn, would suppress the far-right cells, and so on. The resulting activity distribution would be concentrated in the near-right and far-left channels, presumably resulting in the perception of the front surface moving right, the back moving left (Figure 31B). On different trials, activity might instead end up in the near-left and far-right channels,

FIGURE 31

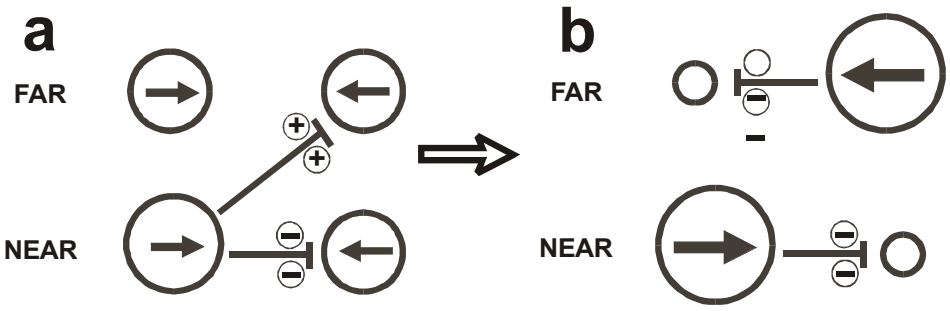

Figure 31: Proposed model which explains how suppression and facilitation in MT could give rise to the illusion of depth. Panel A: A cylinder projection could potentially activate 4 neuronal pools. Panel B: Because of excitatory and inhibitory interactions, activity migrates into opposite-direction, separate-depth channels. This presumably creates the perception of a cylinder rotating in a single direction. Circle diameter denotes magnitude of activity.

depending on the adaptation, or fatigue, of the different channels at the outset of each trial (see also (13)).

In some MT neurons, activity increases when the monkey pays attention to that neuron's preferred direction (10). Assuming our monkeys always attended the cylinder's front surface, this could produce an artifact by increasing activity when the neuron's preferred direction appears in front. However, many neurons responded best when their preferred direction was in back (see above). Moreover, when a neuron preferred a given surface order (based on disparity), it typically responded best when that order was perceived (correlated behavior). This cannot be explained by an attention effect, unless the monkeys learned to selectively attend to one of the two surfaces, depending on the response properties of the neuron currently being tested. This is extremely unlikely.

Area MT is no doubt specialized for motion computation (14, 15), but there is accumulating evidence that it also has a role in three dimensional surface representation. MT neurons have large receptive fields, capable of spatially integrating motion cues; they are direction- and depth-selective, consistent with surface-oriented motion computation; and they exhibit direction-opponency, which may used for surface-specific noise reduction (7, 8, 11, 12, 16). MT neurons are thus well suited to the task of transforming motion cues into information about surfaces and depth. In fact, MT lesions impair monkeys in tasks where three dimensional structure is judged from motion cues (5, 6). Up to now, however, there has been no direct evidence that the perception of structure is linked to activities in MT. Our findings provide this evidence, and as such they suggest that MT has a central role in structure-from-motion perception. Of course, perception may occur in a different area which receives input from MT. But wherever it occurs, our

findings suggest that the perception of structure is ultimately influenced by the segregation of MT activity into separate depth channels.

Methods

Stimuli were shown on a 21 inch CRT display, 57 cm from the monkey's eyes. Isolated neurons were tested to find their approximate receptive field, and subsequent stimuli were centered within this field. Neurons were tested for single-pattern direction selectivity as described previously (8).

Cylinder projections were 7° wide and 7° tall, contained 150 randomly placed dots (8), and rotated at 100°/sec. Cylinders were positioned with their center 3.0-8.2° from the fixation point, and with 10/109 exceptions, no part of the cylinder overlapped the fixation point. All cylinders had their center at zero disparity, so one surface appeared near, the other far, relative to the fixation depth. Monkeys fixated for 0.5 sec before, 1 sec during, and 0.5 sec after the 1 sec cylinder presentation. Selection targets were 5.5° on either side of the cylinder's rotation axis, positioned on a line bisecting the cylinder's height. Monkeys were rewarded with a drop of juice for choosing the target corresponding to the direction of the front surface (for zero-disparity cylinders, rewards were given randomly at a frequency of 80%). Dots were rendered in stereo with an anaglyph system (8).

Monkeys were required to fixate within a 3° square window while the cylinder appeared on the screen. Subsequent analysis showed that eye position remained inside a 1° window in 97% of the trials. The within-trial standard deviation of eye position, sampled at 100 Hz, was 0.05° horizontal, 0.12° vertical.

Most neurons were tested with cylinders containing 5 disparity levels (0, 12.5%, 25%, 50% and 100%) as explained in the text. However, in preliminary tests with 31 of the neurons, only two disparities were tested: one low (0% or 12.5%) and one high (25-100%). Results from these cells were similar to those overall, so they were combined with the remaining cells to form the present data set.

In one of the monkeys, both eye positions were measured simultaneously to calculate the depth of fixation (units of degrees angular disparity). Standard deviation was 0.05° within trials and 0.07° from trial-to-trial. Comparing trials in which opposite surface orders were perceived revealed no differences in fixation depth ($p \geq 0.05$ in 95% of the t-tests; n=53).

Our main analysis involved testing for a response difference ($p<0.05$, t-test) associated with the *perceived* surface order of one or more of the stimuli (each stimulus defined by a given disparity and surface order). Multiple t-testing can in theory increase the false-positive rate, but it cannot account for the high percentage of "correlated" cells (see text), since false-positives have an even chance of being correlated.

Acknowledgments

We are grateful to B. Gillikin and S. Gertmenian for technical assistance and to F. Crick and C. Koch for comments on the manuscript. Supported by grants from the National Eye Institute, the Human Frontier Science Program, and the Sloan Foundation for Theoretical Neurobiology. Address correspondence to Richard A. Andersen, James G. Boswell Professor of Neuroscience, California Institute of Technology, Division of Biology 216-76, Pasadena, CA 91125 (email andersen@vis.caltech.edu).

References

1) B. J. Rogers and M. E. Graham (1982). “Similarities between motion parallax and stereopsis in human depth perception.” Vision Research. 22:261–270.

2) W. R. Miles (1931). “Movement interpretations of the silhouette of a revolving fan.” American Journal of Psychology. 43:392–405.

3) H. Wallach and D. N. O'Connell (1953). “The kinetic depth effect.” Journal of Experimental Psychology. 45:205–217.

4) D. L. Ringach, M. J. Hawken, and R. Shapley (1996). “Binocular eye-movements caused by the perception of three-dimensional structure from motion.” Vision Research. 36:1479–1492.

5) R. M. Siegel and R. A. Andersen (1988). “Perception of three-dimensional structure from motion in monkey and man.” Nature. 331:259-261.

6) R. A. Andersen and R. M. Siegel (1990). “Motion processing in primate cortex.” In *Signal and Sense: Local and Global Order in Perceptual Maps*. Eds. G. M. Edelman, W. e. Gall and W. M. Cowan. Wiley, New York, pp. 163-184.

7) J. H. Maunsell and D. C. Van Essen (1983). “Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity.” Journal of Neurophysiology. 49:1148–1167.

8) D. C. Bradley, N. Qian, and R. A. Andersen (1995). “Integration of motion and stereopsis in middle temporal cortical area of macaques.” Nature. 373:609–611.

9) S. Treue, M. Husain, and R. A. Andersen (1991). “Human perception of structure from motion.” Vision Research. 31:59–75.

10) S. Treue and J. H. R. Maunsell (1996). “Attentional modulation of visual-motion processing in cortical areas MT and MST.” Nature. 382:539–541.

11) R. J. Snowden, S. Treue, R. G. Erickson, and R. A. Andersen (1991). “The response of area MT and V1 neurons to transparent motion.” Journal of Neuroscience. 11:2768–2785.

12) E. C. Hildreth, H. Ando, R. A. Andersen, and S. Treue (1995). “Recovering three-dimensional structure from motion with surface reconstruction.” Vision Research. 35:117–137.

13) M. Nawrot and R. Blake (1991). “A neural network model of kinetic depth.” Visual Neuroscience. 6:219–227.

14) W. T. Newsome, K. H. Britten, and J. A. Movshon (1989). “Neuronal correlates of a perceptual decision.” Nature. 341:52–54.

15) N. K. Logothetis and J. D. Schall (1989). “Neuronal correlates of subjective visual-perception.” Science. 245:761–763.

16) O. Braddick (1993). “Segmentation versus integration in visual-motion processing.” Trends in Neurosciences. 16:263–268.