

Scene Segmentation and Attention in Primate Cortical Areas V1 and V2

DANIEL S. MARCUS AND DAVID C. VAN ESSEN

Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110

Received 6 November 2001; accepted in final form 24 July 2002

Marcus, Daniel S., and David C. Van Essen. Scene segmentation and attention in primate cortical areas V1 and V2. *J Neurophysiol* 88: 2648–2658, 2002; 10.1152/jn.00916.2001. The responses of many neurons in primary visual cortex are modulated by stimuli outside the classical receptive field in ways that may contribute to integrative processes like scene segmentation. To explore this issue, single-unit neuronal responses were recorded in monkey cortical areas V1 and V2 to visual stimuli containing either a figure or a background pattern over the receptive field. Figures were defined either by orientation contrast or by illusory contours. In all conditions, the stimulation over the RF and its nearby surround was identical. Both figure types enhanced the average population response in V1 and V2. For orientation contrast figures, enhancement averaged 50% in V2 and 30% in V1; for illusory contour figures, the enhancement averaged 24% in V2 and 18% in V1. These differences were statistically significant for figure type but not for visual area. In V2, the latency of enhancement to illusory contour-defined figures was longer than that to orientation-defined figures. Neuronal responses were recorded while the monkey performed a directed-attention task. Enhancement to both figure types was observed even when attention was directed away from the figure. Attention slightly enhanced responses in V2, independent of figure type, but did not affect responses in V1. There was no discernible effect of attention on background firing rate in either V1 or V2. These results suggest that scene segmentation is a distributed process, in which neuronal signals at successive stages of the visual hierarchy and over time increasingly reflect the global structure of the image. This process occurs independent of directed visual attention.

INTRODUCTION

The responses of neurons in primary visual cortex (V1) are determined by a combination of factors, including direct activation from the classical receptive field (CRF) and modulatory influences via the nonclassical surround (Allman et al. 1985; Nelson and Frost 1978). The specific effects of the nonclassical surround are highly dependent on stimulus structure. Typically, responses are strongest when the stimulus in the CRF is perceptually salient, as when it “pops out” through local orientation contrast (Knierim and Van Essen 1992) or as part of a global structure such as a contour (Kapadia et al. 1995) or figure (Lamme 1995; Lee et al. 1998; Rossi et al. 2001), suggesting that V1 may participate in integrative processes such as contour integration and scene segmentation. Surround enhancement for figural regions has been observed for figures defined by a number of cues including orientation contrast, disparity, luminance, motion, and color (Zipser et al. 1996) and has been reported to follow the perceptual state of the animal (Lamme et al. 1998; Super et al. 2001).

Address for reprint requests: D. C. Van Essen, Department of Anatomy and Neurobiology, Washington University School of Medicine, 660 S. Euclid, Box 8108, St. Louis, MO 63110 (E-mail: vanessen@v1.wustl.edu).

Despite these advances, a number of issues regarding the role of early visual areas in scene segmentation remain unresolved. First, the degree to which contextual modulation is directly and exclusively implicated in segmentation is unclear. In general, if contextual modulation explicitly reflects object segmentation, it should be absent for nonsegmenting regions and present for all salient figures. The fact that contextual effects have been reported to be sensitive to figure size suggests that this may not be the case (Rossi et al. 2001; Zipser et al. 1996). Alternatively, enhancement in V1 might instead be elicited by stimulus contrast between the CRF and its surround, irrespective of whether the region is part of a figure. Second, it has been unclear whether V1 and V2 play substantially different roles in segmentation. Rossi et al. (2001) reported similar magnitude and spatial extent of contextual modulation in V1 and V2, whereas other studies have reported that V2 is more responsive to surface properties than V1 (Bakin et al. 2000; Zhou et al. 2000). We addressed these issues by recording single-unit responses in V1 and V2 to displays in which the stimulus over the CRF was a textural background pattern, a figure defined by orientation contrast or a figure defined by illusory contours. Our results indicate that the two figural patterns elicit differing degrees of enhancement, and they suggest that V1 and V2 are both involved in segmentation but perhaps to different degrees.

Another unresolved issue is whether “figural enhancement” is dependent on visual attention or whether it occurs in parallel across the visual field. There are plausible grounds for anticipating attentional effects on figural enhancement, given that figure-ground segregation and attention are intertwined processes (Driver et al. 2001) and given that attentional effects have been reported in V1 when monkeys attend to figural patterns (Ito and Gilbert 1999; Roelfsema et al. 1998). We addressed this issue by controlling whether attention was directed toward or away from figural patterns in the cell’s CRF. Our results indicate that figural enhancement is largely independent of where attention is directed.

METHODS

The activity of single neurons in V1 and V2 was recorded from *Dizzy* and *Boozoo*, two male Rhesus monkeys (*Macaca mulatta*) weighing 4.5–7.0 kg. All surgical and experimental procedures conformed to National Institutes of Health and USDA guidelines and were approved by the Animal Studies Committee at Washington University School of Medicine under Protocol 98185. The methods included general procedures described elsewhere (Connor et al. 1997) plus specific procedures summarized in the following text.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Recording and display equipment

Stimulus presentation, experiment procedure, and data acquisition were controlled by customized software. Stimuli were presented on a Silicon Graphics GDM-1640SG RGB monitor driven by a Silicon Graphics 24-bit Indy workstation. The screen was viewed at a distance of 40 cm, subtending $24.2 \times 18.6^\circ$ visual angle. An LCD shutter (NuVision Perceiva 061-0002-00) mounted in front of the monitor alternated between horizontal and vertical polarizing filters synchronously with the 120-Hz monitor refresh (60 Hz for each eye in stereo mode). Static vertical and horizontal polarizing filters were placed in front of the animals' left and right eyes, respectively. The pixel resolution to each eye was $1,280 \times 492$. Stimulus onset was monitored with a photodiode attached to a corner of the display monitor. The photodiode and electrode signals were read from separate channels on the same audio port, and their relative timing was accurate to 0.1 ms.

The monkeys reported discriminations with a custom-built, three-position joystick whose signal was digitized by an eight-channel A/D converter (Immersion). A primate juicer (Crist Instruments) controlled juice rewards. Eye position was monitored using the scleral search coil method (Robinson 1963) with a Rempel Labs EM3 eye movement monitor. The eye movement signal, digitized to eight-bit samples and output at 30 Hz to the computer's serial port, had a resolution of 1.9–2.3 arcmin in the eye coils used.

Neural activity was recorded from cortex using 125- μ m-diam epoxy-coated tungsten electrodes (A-M Systems, No. 575400; initial impedance, 3–5 M Ω) mounted in a guide tube affixed over a 5-mm craniotomy. Signals were amplified by a differential amplifier (A-M Systems), filtered (Krohn-Hite 3700 analog filter), thresholded (Bak DIS-1 window discriminator), and recorded through an audio port channel.

Recording procedures

Recordings were made from areas V1 and V2 of each monkey. Recordings were made from V1 neurons on the opercular surface of the occipital lobe during the early part of each penetration. These neurons were typified by RFs with diameters of 1–1.5 $^\circ$ and eccentricities of up to 8 $^\circ$. Penetrating deeper, recordings were made either from V1 on the posterior bank of the calcarine sulcus or V2 on the posterior bank of the lunate sulcus, depending on the position of the electrode and the folding pattern of the individual monkey's brain. The two areas were easily distinguishable; V1 RFs in the calcarine sulcus had diameters of less than 2 $^\circ$ and eccentricities more than 10 $^\circ$, while V2 RFs in the lunate sulcus had diameters more than 2 $^\circ$ and eccentricities just slightly more eccentric than those of superficial V1 (see Fig. 2).

Once isolated and mapped qualitatively, each unit's spatial sensitivity profile and orientation selectivity were determined quantitatively while the animal performed a passive fixation task. The spatial profile used a 3×7 grid covering 1.5 times the estimated CRF diameter, with bar lengths 0.75 time the estimated CRF diameter; the orientation tuning was based on eight orientations of a bar presented at the quantitatively determined CRF center. The directed-attention test described in the following text was generally run using these quantitatively determined position and orientation preferences. Units preferring horizontal orientations were tested with a slightly oblique orientation that allowed presentation of nonzero horizontal disparity.

The stimuli used to assess segmentation-related activity were full-field line grating patterns that included a circular figural region defined either by orientation contrast (orthogonal disk, Fig. 1A) or by a spatial offset of the grating (parallel disk, Fig. 1B). Whereas the orthogonal disk included both orientation contrast and illusory contour cues, the parallel disk contained *only* the illusory contour cue. The background stimulus (Fig. 1C) included a nonsegmenting grating over the CRF. In all three conditions, the image over the CRF and its near surround was identical.

The line gratings were white (47.0 cd/m 2 at the monitor, 5.7 cd/m 2 measured through the stereo apparatus) on a medium gray background (14.7 cd/m 2 at the monitor, 1.8 cd/m 2 measured through the stereo apparatus). At these luminance levels, visual stimuli appeared bright but cross talk between stereo channels was minimal. Grating line thickness (range: 0.1–0.2 $^\circ$) and spatial frequency (range: 0.75–1.25 $^\circ$) were tuned to maximally stimulate each isolated neurons. Figural patterns were generally displayed at 2.5 times the CRF diameter, but for a few cells were reduced to insure that the figure border was at least 0.5 $^\circ$ from the fixation point.

Behavioral paradigm

In the directed-attention task (Fig. 1D), the monkey was trained to report the direction of a stereoscopic depth change within a cued figure while ignoring distracters. Each trial began with the monkey fixating a small black disk (0.1 $^\circ$ across). Throughout the trial, fixation was maintained within a small window (0.75 and 0.9 $^\circ$ diameter for *Dizzy* and *Boozoo*, respectively.) After 500 ms, a colored disk centered on the fixation point appeared for 500 ms, with the color indicating the side of the screen containing the behaviorally relevant figure in the upcoming stimulus (red for left, green for right). After a brief hold period (500 ms for *Dizzy*, 250 ms for *Boozoo*), the stimulus was displayed. The stimulus consisted of a figure in the hemifield opposite the CRF and (except in the background condition) a figure centered over the CRF. To ensure that the monkey attended to the cued figure, he was required to report the direction of a stereoscopic depth change in the cued figure while ignoring depth changes in the noncued figure and background grating. Depth changes appeared as stereoscopic Gaussian bumps and dents with a diameter 0.75 times that of the figure. The depth of the bump or dent was adjusted to near the animal's behavioral threshold; the peak disparity was typically 0.25–0.4 $^\circ$, depending on eccentricity and daily fluctuations in performance levels. Depth changes occurred with equal independent likelihood at the cued and noncued positions and could occur at any time from 0 to 1,000 ms after stimulus onset. Trials in which depth changes occurred earlier than 500 ms were excluded from data analyses. Each depth change was masked 500 ms after appearing. The monkey received a water reward if he correctly reported the direction of the depth change within the cued figure within 800 ms of the onset of the depth change.

Seven repetitions of each stimulus condition plus catch trials in which depth changes occurred earlier than 500 ms resulted in a total of 49 trials for the behavioral test. Trials were randomized within blocks that contained a single repetition of each condition. After an incorrect trial, a 3- to 5-s pause was included prior to the subsequent trial to punish the monkey. Performance on the task was consistently better than 80%. Incorrect trials included a comparable incidence of eye position and discrimination errors.

The monkeys' ability to discriminate rapidly-masked depth changes within a cued figure while ignoring distracter depth changes in the nearby surround and in a noncued figure indicates that they selectively attended to the appropriate figure. Because the depth changes occurred with a variable onset the monkeys also had to remain vigilant throughout the trial. Additionally, because the cue indicated only the hemifield in which the relevant figure was located, the task forced the monkeys to process the shape of the figures prior to reporting a discrimination. Both monkeys were able to locate figures at random positions on the cued side of the screen, indicating that figural segmentation occurred even at moderate eccentricities.

Two different layouts were used for the background condition, which was interleaved with the figural conditions. For *Dizzy*, a figure was included in the CRF hemifield but was positioned at least 2.5 CRF diameters away from the CRF center. For *Boozoo*, no figure was included in the CRF hemifield. It is unlikely that these differences substantially affected responses because responses to background

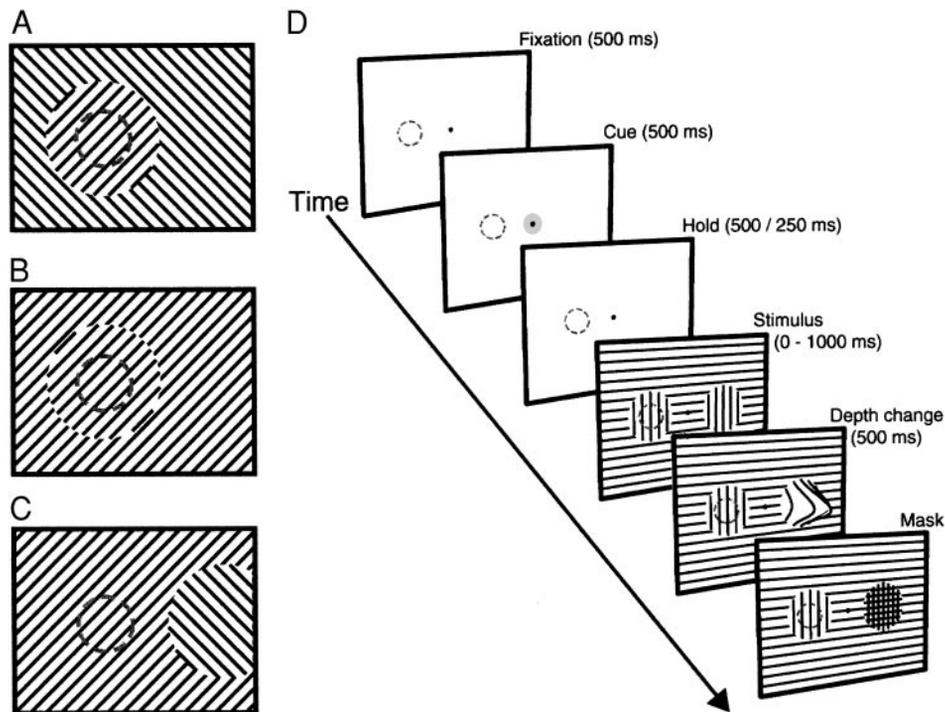


FIG. 1. Stimuli and behavioral paradigm. A typical receptive field (RF) is schematized as a dashed gray ring. The orientation and spatial frequency of the grating over the RF was tuned to the preferred values for the individual units. In all conditions, the background grating extended over the entire screen. *A*: in the orthogonal disk condition, orientation contrast and illusory contour cues both contribute to the segregation of the figure. *B*: in the parallel disk condition, only illusory contour cues contribute to the segregation of *C*. In the background condition, the segregating figure is positioned distant to the RF. A crucial feature of the stimulus set is that the image within the RF and its nearby surround is identical in all 3 conditions. *D*: each frame represents the indicated epoch in a typical trial. Trials began with the monkey fixating a small disk at the center of the screen for 500 ms. A color cue then appeared for 500 ms around the fixation point, indicating the side of the screen on which the monkey was to attend to a subsequent figure. A green cue, shown here as a light gray disk, indicated the right side of the screen, while a red cue indicated the left side. After a hold period (500 ms for 1 animal, 250 for the other), the stimulus appeared. The stimulus consisted of a background grating and a figure on each side of fixation. By manipulating the relative position and orientation of the figures and background grating, we were able to observe responses to each of our stimulus types under attending-to and -away conditions. In this example, the monkey was cued to attend to the figure on the right side of the screen, away from the RF. To ensure that the animal attended as cued, he was required to report the direction of a depth change in the cued figure while ignoring depth changes in the noncued figure and background grating. Here, the monkey would pull the joystick to report a "bump" to receive a juice reward. Depth changes could occur at any time during the 1,000-ms stimulus period. A mask appeared over depth changes after 500 ms. The masks consisted of lines at random depths at parallel and orthogonal orientations.

stimuli are reportedly unaffected by proximity or complete absence of a figure (Lamme 1995; Lamme et al. 1999).

Data analysis

Raw firing rates were determined over the period between 40 and 500 ms after stimulus onset. This period begins with the average onset latency observed in V1 and V2 and proceeds through the minimum stimulus duration. Evoked responses were determined by subtracting the spontaneous baseline activity, calculated over the 250 ms immediately preceding the cue on a trial-by-trial basis.

To measure the effect of figure type on the evoked responses, we used a figural modulation index (FMI). The FMI represents the difference between the average responses to the figure and background conditions normalized by the average response to the background. It was calculated as follows

$$\text{FMI} = (R_{\text{figure}} - R_{\text{background}}) / (R_{\text{background}} + \text{STD})$$

where R was usually the evoked response for the given condition. For the few cells in which the evoked response to the background condition was negative, the raw firing rate was used for each condition. By including the STD (the averaged SD for the figure and background conditions) in the denominator, the effects of response

variance were minimized. All analyses using the FMI were also conducted using indices that excluded the STD factor. These analyses yielded similar but more scattered results. Separate FMI values were calculated for the parallel disk and orthogonal disk figure types. To measure the effect of attention, we calculated an attentional modulation index as follows

$$\text{AMI} = (R_{\text{to}} - R_{\text{away}}) / R_{\text{away}}$$

Again, evoked responses were generally used, but the raw firing rates were used if the response in the attending-to condition was negative.

To ensure that the monkey's attention was consistently directed to figures, not just locations in space, the behavioral paradigm did not include a condition in which the monkey attended to the background stimulus. Consequently, analyses of the effects of figure type on neuronal responses included only trials in which the animal was attending *away* from the stimulus over the CRF. Tests for interactions between figure type and attention included only the parallel disk and orthogonal disk stimuli. In several analyses, when there was no significant interaction between figure type and attention, trials with the same attentional condition but different figural condition were collapsed into the same group, to improve signal to noise.

Data analysis was conducted in Matlab, using commercial and

custom scripts. In cases where the measurements and indices were not normally distributed (as determined by Lilliefors test), nonparametric versions of tests were used in place of the standard tests, including Student's *t*-test and one- and two-way ANOVAs.

RESULTS

A total of 51 V1 and 69 V2 well-isolated single units are included in our data set (Fig. 2A). From *Dizzy*, 29 V1 and 45 V2 units were recorded from three craniotomies in the left hemisphere; CRF eccentricities ranged from 2 to 8°. From *Boozoo*, 22 V1 and 24 V2 units were recorded from three craniotomies in the right hemisphere; CRF eccentricities ranged from 3 to 9°. Overall, CRF sizes ranged from about 0.5–4° in diameter (Fig. 2B).

During manual plotting of response properties of isolated

cells, about 20–30% of the cells gave little or no response to grating stimuli, presumably because of strong end- and/or side-inhibition. For cells that were poorly responsive to the test stimuli, the main test was not run or was discontinued partway through. Most of the remaining cells gave a diminished response to the grating stimuli relative to the most effective bar stimulus, indicating that responses in the behavioral task were well below saturation.

Effects of figure type in V1 and V2

An example V1 unit's response to each of our test conditions is shown in Fig. 3, A and B. The raster plot (Fig. 3A) shows the spikes over the entire course of each trial grouped by stimulus type and attentional condition. Each trial is aligned to the stimulus onset, indicated by the blue vertical line. The moderate spontaneous activity observed in this unit during fixation period was unaltered by the appearance of the color cue around fixation. Spontaneous activity during the hold period was unaffected by direction of attention, as indicated by similar spiking patterns between the green and red clusters, and was similar to that during the fixation interval. A brisk, sustained response began shortly after stimulus onset for all of the test conditions. A number of differences between conditions are evident from visual inspection of the rasters and are quantified for each condition in the plots of mean evoked response (Fig. 3B). The mean evoked response to the orthogonal disk was significantly greater than when the parallel disk or background stimulus was presented (one-tailed *t*-test, $P < 0.05$). The response to the parallel disk was slightly but not significantly greater than to the background stimulus. For both the parallel and orthogonal disk stimuli, the responses in attending-to and attending-away conditions were not significantly different (*t*-test, $P > 0.05$).

An example V2 unit's response to each of the test conditions is shown in Fig. 3, C and D. The raster plots (Fig. 3C) reveal low spontaneous activity prior to stimulus onset that was not modulated by attentional condition. Shortly after stimulus onset, the unit fired strongly to both figure types but only moderately to the background stimulus. The strong responses to figures were sustained over the entire stimulus presentation.

Looking at the mean evoked responses for each condition (Fig. 3D), both orthogonal and parallel disk stimuli elicited about twice the response as the background stimulus. This difference was significant for both figure types (1-tailed *t*-test, $P < 0.01$). The mean firing rates were also slightly higher in attending-to conditions than attending-away conditions. This effect was not significant for either figure type in this unit.

To assess how figure type affected the populations in V1 and V2 and in each animal, the mean population firing rate to each figure condition was calculated for each animal. Figure 4 shows the average firing rates (A) and the average normalized firing rates (B) to the three figural conditions for the V1 populations from each animal, using data only from attending-away trials (see METHODS). In these analyses, raw firing rates were used to circumvent occasional negative values found with evoked responses. The plots indicate that the population from *Dizzy* had a higher mean firing rate than the population from *Boozoo* for each condition. However, the general pattern of activity for each condition was similar between animals. The plots of the normalized rates (calculated as the average within-

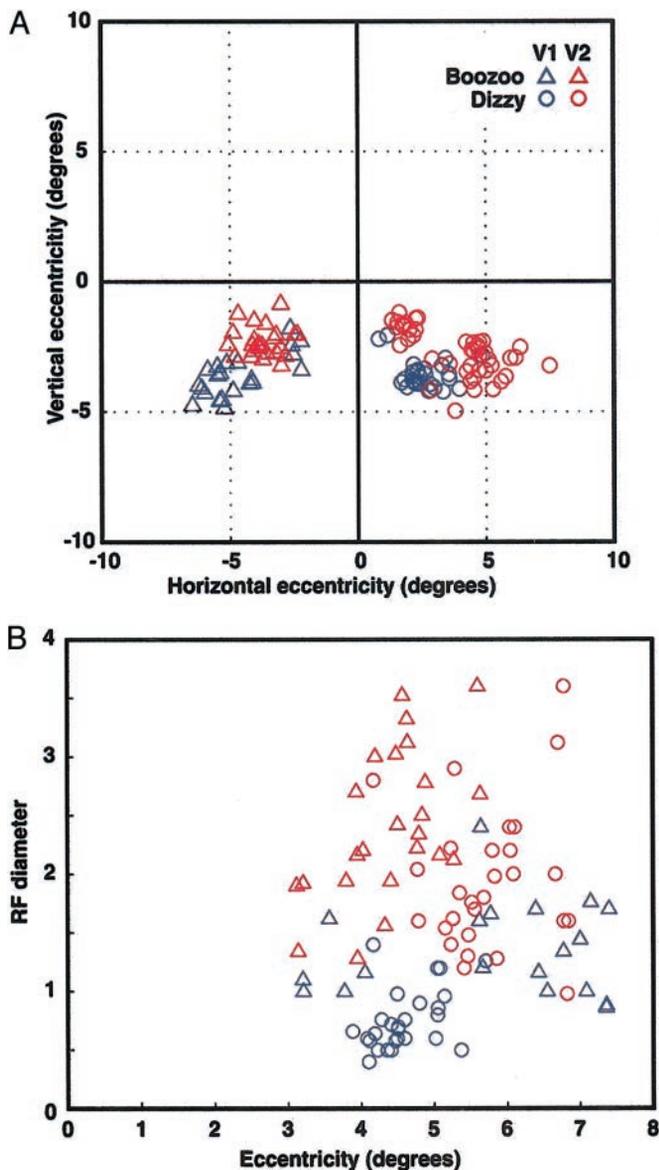


FIG. 2. Data set. A: RF positions for V1 and V2 of each subject. From *Dizzy*, 29 and 45 units were recorded from V1 and V2, respectively. From *Boozoo*, 22 and 24 units were recorded from V1 and V2, respectively. B: estimated RF diameters for the majority of the V1 and V2 cells recorded from *Boozoo* and *Dizzy*.

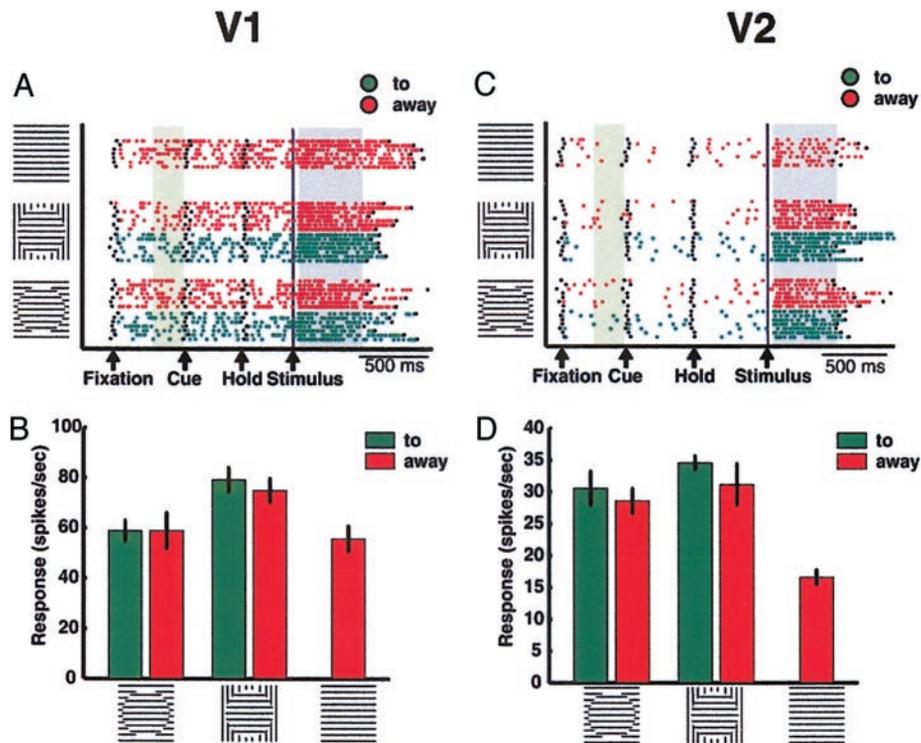


FIG. 3. Example neurons. *A* and *B*: example V1 unit, *B8.15.1*. *A*: rasters are grouped by stimulus type, as indicated along the vertical axis, and attentional condition as indicated by color. Trials are aligned by stimulus onset, represented by the blue vertical line. Additional trial events are labeled along the horizontal axis. The response to the orthogonal disk is stronger than that to the parallel disk or background stimulus beginning at the earliest part of the response and continuing throughout the stimulus period. *B*: mean response by figure type and attentional condition. To calculate the mean response, the total firing rate is taken for each trial over the interval shaded in the raster plot (40–500 ms after stimulus onset). The spontaneous rate is then calculated for that trial over the 250 ms immediately preceding the behavioral cue and subtracted from the total firing rate to obtain the mean response for each trial. The mean response, as plotted here and used throughout the paper, represents the average trial-wise mean response for a given condition. In this example, the mean response to the orthogonal disk is significantly greater than that to the background and to the parallel disk (1-tailed t -test, $P < 0.05$), while the response to the parallel disk is not significantly different from the response to the background stimulus. No significant effect of attention is apparent. *C* and *D*: example V2 unit, *D7.3.1*. *C*: the sustained response beginning shortly after stimulus onset is much stronger for both figural stimuli than for the background stimulus. *D*: a significant enhancement is observed for both figure types relative to background ($P < 0.05$). Additionally, modest enhancement is observed in the attending-to conditions.

cell mean firing rate to each condition normalized by the within-cell mean firing rate to the background condition) confirm this point. In a two-way ANOVA on the normalized responses using figure type and subject as factors, only the figure type factor was significant ($F = 17.01$, $P < 0.00001$), and no interaction was apparent between the factors. Collapsing across subjects, average population activity to the orthogonal disk was enhanced by 30% relative to the background condition (Bonferroni t -test; $P < 0.01$), whereas average population activity to the parallel disk was enhanced by 18% (Bonferroni t -test; $P = 0.01$). Additionally, the difference between the average rate for the orthogonal disk and parallel disk was significant (Bonferroni t -test; $P < 0.05$).

Figure 4, *C* and *D*, shows the average firing rates and average normalized firing rates for the V2 population in each animal. In a two-way ANOVA using figure type and subject as factor, only the figure type factor was significant ($F = 9.48$, $P < 0.0001$) and no interaction was apparent between the factors. Collapsing across subjects, the average normalized firing rate was enhanced by 50% to the orthogonal disk (Bonferroni t -test; $P < 0.01$) and by 24% to the parallel disk (Bonferroni t -test; $P < 0.05$).

The preceding analyses indicate that in both V1 and V2, the

populations as a whole were significantly enhanced by both figure types, with the orthogonal disk leading to significantly stronger modulation than the parallel disk. To determine how figure type impacted the responses of individual cells, within-cell ANOVAs were run on each unit. In V1, 14 of 51 units (27%) were significant for figure type ($P < 0.05$). Post hoc pair-wise comparisons of the figural responses of the units that were significant in the ANOVAs indicated that all 14 had a significantly stronger response to orthogonal disks than to the background condition, whereas only two had a stronger response to the parallel disk than to the background condition (Tukey's LSD; $P < 0.05$). Additionally, six units had a significantly stronger response to the orthogonal disk than to the parallel disk. In V2, 27 of 69 units (39%) were significant in the within-cell ANOVAs on stimulus type ($P < 0.05$). In post hoc tests on the significant units, most (24) had significantly stronger responses to the orthogonal disk than to the background stimulus, and eight had significantly stronger responses to the parallel disk than to the background stimulus. More than half (20) had a significantly stronger response to the orthogonal disk than to the parallel disk.

While the population analyses indicate that both figure types significantly enhanced responses in both areas, the within-cell

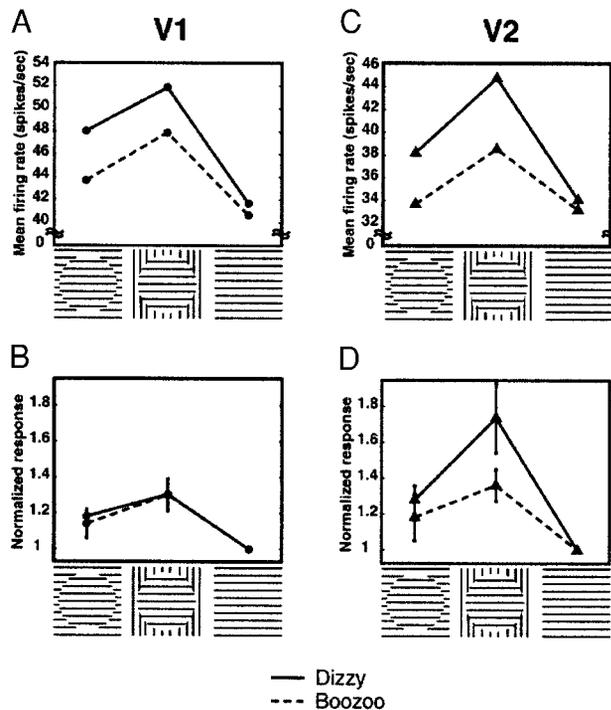


FIG. 4. Population response to figural stimuli. Each subject is plotted separately. The response was averaged between 40 and 500 ms after stimulus onset. Only attending-away trials are included for each figural condition. *A*: The raw responses in V1 were somewhat stronger for *Dizzy* than *Boozyoo* but show a similar pattern of enhancement to figural stimuli. *B*: to compare effects between animals, the responses in each condition were normalized within-cell by the response to the background stimulus. The normalized responses in V1 were very similar between animals. The average response to the orthogonal disk was 30% higher than that to the background, while the average response to the parallel disk was 18% higher than that to background. *C*: as in V1, the raw response in V2 to each figure type was stronger in *Dizzy* than in *Boozyoo*. *D*: the normalized responses in V2 show strong enhancement to both figure types, relative to the background stimulus (50% for the orthogonal disk, 24% for the parallel disk). Enhancement to the orthogonal disk appears particularly strong in *Dizzy*.

analyses indicate that only a minority of the cells in each area was significantly modulated. Indeed, in V1 the parallel disk led to a significant effect in fewer than 4% of the cells. To get a better sense of the overall distribution of figural effects, figural modulation indices (FMIs) were calculated for each cell (see METHODS). The scatter plots in Fig. 5 illustrate the parallel and orthogonal FMIs for each unit in V1 (*A*) and V2 (*B*), respectively. Units that were significant ($P < 0.05$) in the preceding ANOVA on stimulus type are plotted as filled markers. Most units were more strongly enhanced by the orthogonal disk than the parallel disk, as indicated by their position above the unity line on the plots. The distance of each unit from the diagonal is indicated in the histogram plotted along the top right of each scatter plot. In both areas, the distribution mean (0.09 for V1, 0.15 for V2) is shifted significantly above the diagonal (t -test, $P < 0.0001$), indicating that on average the magnitude of enhancement was greater for orthogonal disks than parallel disks. In V1, all of the significant units were more strongly enhanced by orthogonal disks than parallel disks. In V2, the enhancement for parallel disks exceeded that for orthogonal disks in 4 units, but the majority showed stronger enhancement to the orthogonal disks.

The FMIs were also used to examine whether modulation

differed significantly by cortical area. A two-way ANOVA was conducted using the FMI and cortical area as factors. This analysis revealed a significant effect for figure type ($P < 0.0001$) but not for cortical area ($P = 0.1401$). No interaction was found between factors.

Inspection of the scatter plots and histograms reveals that the distributions of FMIs were fairly continuous in both V1 and V2 and show no evidence for a specific subset of cells dedicated to figure-ground processing. Although histological reconstruction of electrode tracks was not feasible, it was clear from our electrode depth readings that units with strong figural enhancement were found in both superficial and deep cortical laminae of both V1 and V2. Penetrations were generally aimed perpendicular to the cortical surface, and within a single penetration, we not uncommonly observed units with strong enhancement effects throughout the cortical depth. Modulation strength was also not dependent on the relative size of a unit's RF, in that strong effects were found for units with relatively small and relatively large RFs for a given eccentricity.

Time course of responses to figural stimuli

Figure 6 shows the average population responses over time to the different stimulus types for V1 and V2. The population responses in V1 (Fig. 6, *A* and *B*) show a pronounced onset transient, peaking approximately 40 ms after stimulus onset followed by weaker sustained activity. As expected from the FMI analyses, the response to the parallel disk was modestly stronger than that to the background stimulus (Fig. 6*A*). A trend toward enhancement was present early in the response, with a single significant bin appearing 110 ms after stimulus onset, but consistent enhancement was evident only in the late part of the response. For the orthogonal disk (Fig. 6*B*), enhancement reached significance at 80 ms and was sustained throughout stimulus presentation.

Figure 6, *D* and *E*, shows the average time course of the population responses to the different figural conditions for the entire V2 population. The response patterns are similar to those in V1 in latency and time of initial peak. As in V1, enhancement to the orthogonal disk (Fig. 6*E*) began early and was sustained throughout. Enhancement to the parallel disk was smaller in magnitude and began only after a delay of around 150 ms (Fig. 6*D*).

To improve the sensitivity for detecting temporal differences, the figural modulation index for each cell was calculated during selected epochs of the response. We chose four epochs: onset peak (40–80 ms), postpeak (80–120 ms), intermediate (120–200 ms), and sustained (200–500 ms) but emphasize that these epochs are not meant to reflect natural temporal subdivisions.

Figure 6*C* shows the mean FMI for the V1 population for each figure type during each epoch. The mean enhancement changed little for either the parallel disk or orthogonal disk during the course of the response. During each period, most units showed no modulation or slight enhancement, but even during the earliest epoch a handful of units showed substantial enhancement. The mean FMI for the parallel disk was significantly greater than zero in every epoch except the postpeak period (1-way Student's t -test, $P < 0.05$), and the mean FMI for the orthogonal disk was significantly greater than zero during every epoch ($P < 0.05$). To test whether FMIs were

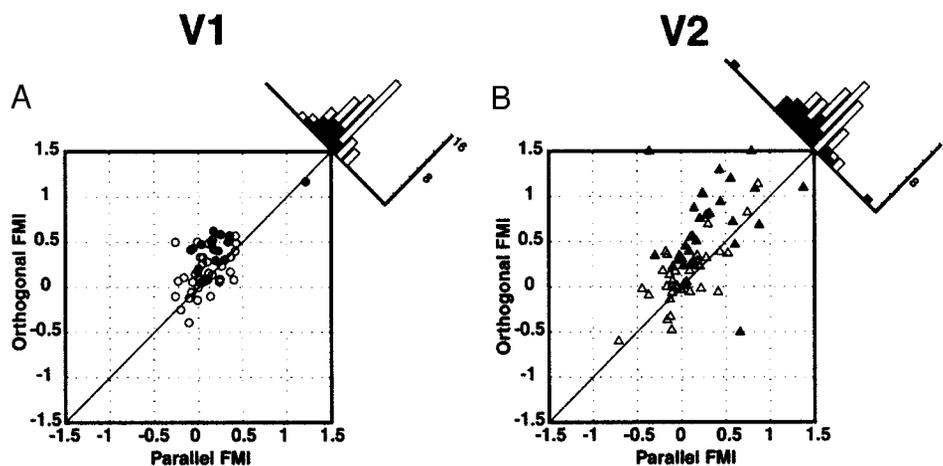


FIG. 5. Scatter plots of figural modulation indices. The plots include data from both subjects. Each marker represents a single unit. Units that were significant in an ANOVA on figure type are plotted as filled markers. The histograms at the top right of each plot indicate the number of units at the indicated distance from the unity line. *A: V1.* Most units lie in the top right quadrant, indicating that their responses were stronger to both figure types than to the background. The mean figural modulation index (FMI) in V1 for the parallel disk was significantly greater than 0 for *Dizzy* (1-tailed Student's *t*-test; $P = 0.0004$) but not *Boozoo* ($P = 0.21$). The mean FMI for the orthogonal disk was significantly greater than zero for both animals (*Dizzy*: $P = 0.00001$; *Boozoo*: $P = 0.0003$). The distribution is shifted significantly above the unity line, with a mean distance of 0.09 (*t*-test, $P < 0.00001$), indicating that on average responses in V1 were more strongly enhanced by the orthogonal disk than the parallel disk. A 2-way ANOVA on the FMIs using figure type and subject as factors revealed a significant difference between FMIs ($F = 7.44$, $P = 0.008$) as well as a significant difference between subjects ($F = 4.39$, $P = 0.039$), with no interaction between the factors. *B: V2.* As in V1, most units lie in the top right quadrant and are above the unity line. The mean FMI for the parallel disk was significantly greater than 0 for both animals (1-tailed Student's *t*-test; *Dizzy*: $P = 0.003$; *Boozoo*: $P = 0.02$). The mean FMI for the orthogonal disk was significantly greater than 0 for both animals (*Dizzy*: $P = 0.000002$; *Boozoo*: $P = 0.0006$). The distribution is shifted significantly above the unity line with a mean distance of 0.15 ($P < 0.000001$). In a 2-way ANOVA, a significant effect of figure type was observed ($F = 7.81$, $P = 0.006$) with no significant effect of subject ($F = 0.44$, $P = 0.51$) and no interaction between the factors.

different during different epochs, a two-way ANOVA was conducted using epoch and figure type as factors. The effect of figure type was significant ($P < 0.0001$) while that of responses epoch was not.

Figure 6*F* shows the mean FMI for the V2 population during the same epochs as in V1. The strength of enhancement in V2 varied over time somewhat more than in V1. For the orthogonal disk, the mean FMI was significantly greater than zero during each epoch (Student's *t*-test, $P < 0.05$) and increased over the course of the stimulus display. For the parallel disk, the mean FMI significantly exceeded zero only for the intermediate (120–200 ms) and sustained (200–500 ms) epochs. In a two-way ANOVA, we observed significant effects of figure type ($P < 0.0001$) and response epoch ($P = 0.041$). However, post hoc comparisons did not indicate a significant difference between any of the epochs for the parallel or orthogonal disk. The same ANOVA test was run on the subset of cells ($n = 27$) that showed significant figural modulation (i.e., those shown as filled circles in Fig. 5*b*). Here, both figure type and epoch were highly significant ($P < 0.01$). Post hoc comparisons demonstrated a significant difference between the first and last epochs of the response to the orthogonal disk but did not reveal a clear onset of modulation.

Effects of attention in V1 and V2

The behavioral paradigm was designed to reveal whether attention affects responses to figures in V1 and V2 and whether any effects differ between areas or between figure types. To address these issues for the population as a whole, the average population firing rates during presentation of figural stimuli were calculated for the attentional conditions, after collapsing

across figural conditions, for V1 (Fig. 7*A*) and V2 (Fig. 7*D*). Responses were then normalized by the mean response to the attending-away condition (Fig. 7, *B* and *E*). While the raw population firing rate in V1 (Fig. 7*A*) for the units from *Dizzy* were slightly higher than those from *Boozoo*, the normalized plot (Fig. 7*B*) suggests at most a slight (5%) enhancement for attending-to versus attending-away. However, in a three-way ANOVA using factors of attentional condition, figural condition, and subject, no significant effects of attention were observed and there were no significant interactions between factors.

In V2 (Fig. 7, *D* and *E*), the raw firing rates (*D*) were again somewhat stronger for *Dizzy* than *Boozoo*, but the normalized rates were very similar (*E*). Responses were stronger by 7% in the attending-to versus attending-away condition. In a three-way ANOVA as in the preceding text, the attentional effect was highly significant ($F = 7.01$, $P < 0.01$). No interaction was observed between figural condition and attentional condition or subject and attentional condition.

The requirement that the monkeys attend to diametrically opposed figures in the attending-to and attending-away conditions led to a slight bias in the average fixation position toward the attended figure. The average distance between mean fixation in attending-to and -away trials was 4.8 arcmin, which was statistically significant (ranksum test, $P < 0.01$). While the average eye position was significantly biased toward the attended figure, these biases were unlikely to systematically enhance responses because the figural pattern covered the entire CRF and extended well beyond. A more likely effect of eye position biases would be to add variance to the neuronal responses, particularly in V1, where CRFs are smaller (Fig.

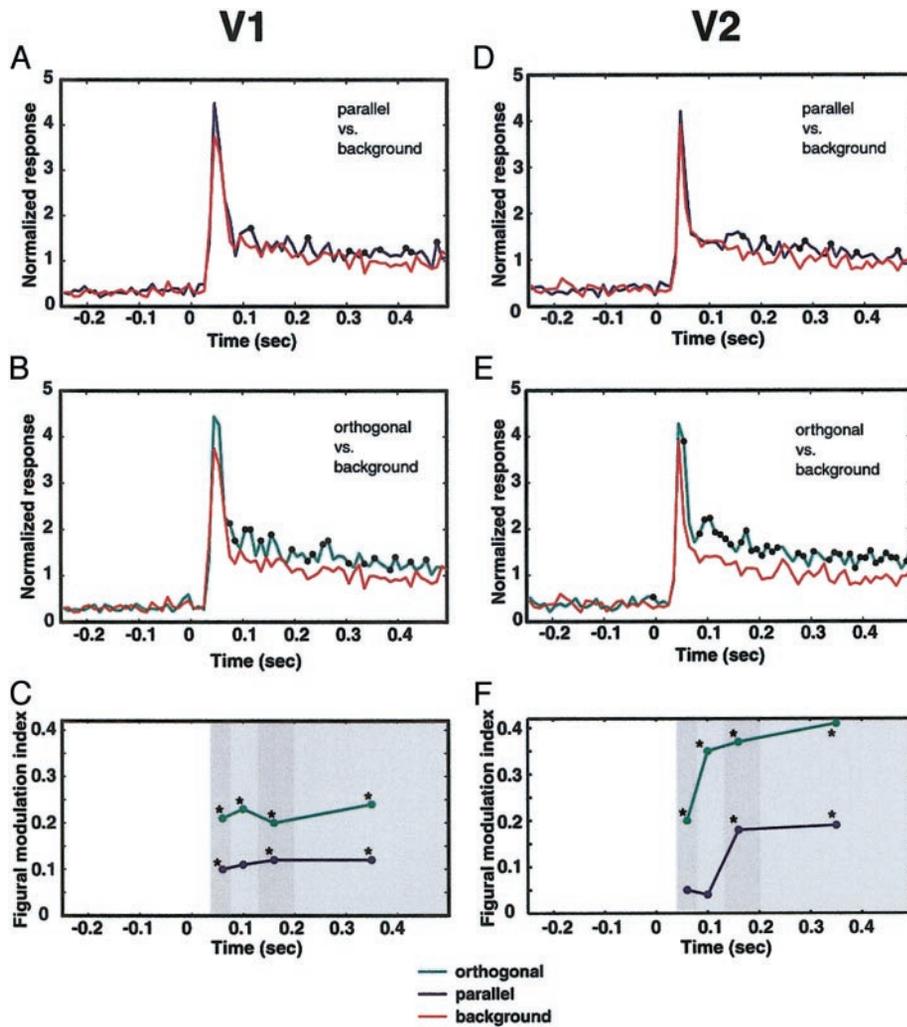


FIG. 6. Time course of figural effects. The spikes from each unit were first segregated into 10-ms bins, aligned to stimulus onset at *time 0*. Each bin was then normalized to the overall mean firing rate for all attending-away conditions for that unit. For each condition, each bin was averaged for the entire population, as plotted here. Bins that are significantly greater for the figural condition than the background condition are indicated with black dots (1-tailed *t*-test, $P < 0.05$). *A*: the response to the parallel disk in V1 shows some sporadic enhancement throughout the stimulus period. Most of the significant bins are clustered in the later part of the response. *B*: the response to the orthogonal disk in V1 shows steady enhancement throughout the stimulus period. *C*: mean AMI values in V1 during each of the epochs (indicated by alternating shaded regions) are plotted for the parallel (blue) and orthogonal (green) disks. Asterisks indicate epochs during which the mean AMI significantly exceeded 0. *D*: enhancement to the parallel disk in V2 begins only after a substantial delay, after which enhancement is modest but steady. *E*: enhancement to the orthogonal disk in V2 begins early in the response and is strong throughout the stimulus period. *F*: mean AMI values in V2 during each epoch.

2*B*). This may have contributed to the larger standard error in normalized responses in V1 (Fig. 7*B*) versus V2 (Fig. 7*E*).

As with the figural conditions, attentional effects were analyzed at the level of individual neurons as well as the populations as a whole. To begin, a two-way ANOVA was conducted on the evoked responses in each cell using location of attention (to vs. away) and figure type (orthogonal disk vs. parallel disk) as factors. The behavioral paradigm did not include conditions in which the animal attended to a background grating over the CRF, so the background stimulus was not included in the ANOVA. In V1, only 3 of 51 cells were significant for attention ($P < 0.05$). Although this number would be expected by chance, all had significantly stronger responses when attention was directed over the CRF (post hoc Bonferroni *t*-test, $P < 0.05$). As expected from the previous analyses of figural effects, a moderate number of units (10 of 51) were significant for figure type. In V2, 5 of 69 cells were significant for attention, and in each of these units, the preference was for the attending-to condition (1-tailed *t*-test, $P > 0.05$). Many units (22 of 69) were significant for figure type.

An attentional modulation index (AMI) was used to ascertain the range of effects within each population. Figure 7*C* shows the distribution AMIs in V1. The mean of this distribution (0.06) was not significantly greater than zero (1-tailed *t*-test, $P > 0.05$). Figure 7*F* shows the distribution of AMIs in

V2. The mean of this distribution was 0.07, which significantly exceeded zero (1-tailed *t*-test, $P = 0.003$).

Although interactions between attention and figure type were rare in the two-way ANOVA tests, it remains possible that on average, attention had a greater effect on one figure type than the other. Accordingly, we calculated an AMI for each cell for each figure type. For V1, the average AMI was 0.016 for the parallel disk and 0.14 for the orthogonal disk; for V2, the average AMI was 0.077 for the parallel disk and 0.069 for the orthogonal disk. These differences were not significant for either area (ranksum test, $P = 0.29$ for V1, $P = 0.83$ for V2).

The population time course of responses under the attending-to and -away conditions is shown for the V1 in Fig. 8*A* and for V2 in *B*. As expected from the mean response analyses, little difference was apparent between the two conditions in V1. In V2, the response was somewhat stronger in the attending-to than in the attending-away condition. This enhancement was only apparent in the later part of the response, starting approximately 300 ms after stimulus onset. To further examine the effects of attention during the later part of the response, the AMI was calculated after collapsing the figural conditions for each cell in V2 over the period of 300–500 ms after stimulus onset. The mean AMI during this period was 0.28 for *Dizzy* (12/45 cells with $P < 0.05$) and 0.073 for *Boozoo* (2/24 cells

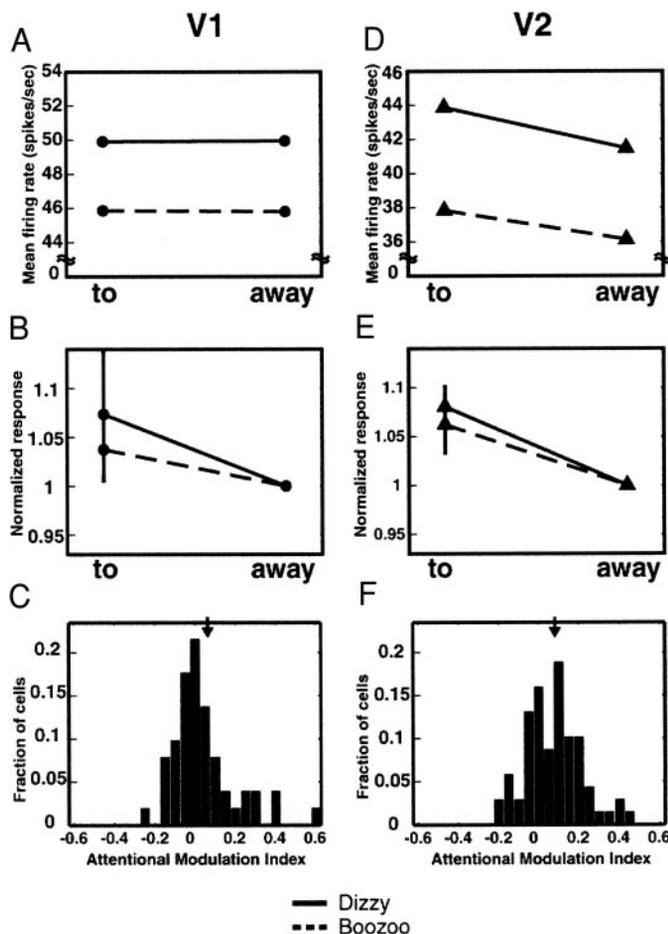


FIG. 7. Population firing rate to attentional conditions. Each subject is plotted separately. Responses are taken over the 1st 500 ms of stimulus presentation and are collapsed across figural conditions. *A*: the raw responses in V1 indicate little or no effect of attentional condition in either animal. *B*: the normalized responses in V1 suggest a slight enhancement (approximately 5%) when attention is directed to a figure over the RF. The apparent discrepancy between the raw and normalized responses can be accounted for by the fact that cells with stronger firing rates bias the raw responses. *C*: attentional modulation index (AMI) for V1 population. The mean of the distribution (indicated by the \downarrow) is 0.06 and is not significantly greater than 0 (1-tailed t -test, $P > 0.05$). *D*: the raw responses in V2 were stronger in *Dizzy* than in *Boozoo*, but both animals showed a similar increase in response in the attending-to condition. *E*: the normalized responses in V2 are similar between animals and suggest about a 7–8% enhancement in the attending-to condition. *F*: the mean AMI is 0.07 and is significantly greater than 0 ($P = 0.003$). While most units were enhanced by attention, few units were enhanced by more than 20%.

with $P < 0.05$). The majority of the effect during the later period thus seems to be attributable to one of the monkeys.

Attention has been reported to increase background firing rates modestly under certain circumstances (Chelazzi et al. 1993; Haenny and Schiller 1988; Luck et al. 1997). In our task, the animal was cued to the left or right hemifield and then held fixation for 0.5 s before the stimulus appeared. During this period, the animal might have directed his attention to the hemifield to which he was cued or even to an expected position within that hemifield. Accordingly, we calculated the average background activity during the 250 ms prior to stimulus onset for each unit in attending to and attending away conditions. There was no significant difference in V1 (10.3 vs. 11.1 spikes/s for attending-to conditions vs. attending-away condi-

tions; $P = 0.78$ by a 2-tailed t -test) or in V2 (10.4 vs. 10.6 spikes/s in attending-to conditions vs. attending-away conditions; $P = 0.93$).

DISCUSSION

The modulatory effects demonstrated in this study indicate that responses in both V1 and V2 are enhanced by figures defined by a variety of cues. There was a trend toward increased magnitude and frequency of effects in V2 relative to V1, but this was not statistically significant. Thus our results suggest that segmentation cues are captured over multiple stages of processing and perhaps to a greater degree in higher areas. That these effects are apparent even when attention is directed to a distant location indicates that these aspects of segmentation operate preattentively and in parallel across the retinal image.

Figural modulation in V1

Several studies have previously reported enhancement in V1 to figures that, like our orthogonal disk, were defined by orientation contrast (Lamme 1995; Lee et al. 1998; Super et al. 2001; Zipser et al. 1996). In comparison to our results, Zipser et al. (1996) reported stronger average modulation with textured figures in V1 (69% vs. our 30%) and with greater frequency (significant enhancement in 52% of units vs. our 27%). In contrast, Rossi et al. (2001) reported substantially less modulation than reported here using stimuli nearly identical to those used by Zipser et al. (1996). Rossi et al. (2001) suggested that attentional effects related to the saccade task used by Zipser et al. (1996) may have led to enhancement unrelated to contextual effects. This seems unlikely, though, because the effects we observed were identical regardless of where attention was directed (see also Lamme et al. 1998).

Regardless of its exact magnitude, several observations suggest that modulation in V1 to orientation contrast-defined figures does not represent a general segmentation or figure-ground segregation mechanism: it is highly dependent on figure size (Lee et al. 1998; Rossi et al. 2001; Zipser et al. 1996); it is apparent for orientation contrast regions that are not completely bounded and do not segregate (Rossi et al. 2001); and, conversely, it is weak for the parallel disks used here, which do segregate. Instead contextual modulation in V1 seems primarily involved in identifying local regions that contrast with their surround in one or more low-level cues, including orientation, luminance, disparity, or motion. This activity likely represents an early stage of a distributed segmentation process.

Segmentation and surface representation in V2

The enhancement for orientation contrast figures tended to be greater in magnitude and frequency of occurrence in V2 than in V1 in our data set. While others have found little difference in the nature of effects between V1 and V2 (Rossi et al. 2001), our findings are consistent with a number of reports that V2 is more closely involved in boundary and surface processes than V1 (Bakin et al. 2000; Lee and Nguyen, 2001; von der Heydt et al. 2000; Zhou et al. 2000).

A number of these reports indicate that V2 plays a critical role in boundary processes such as amodal completion (Bakin

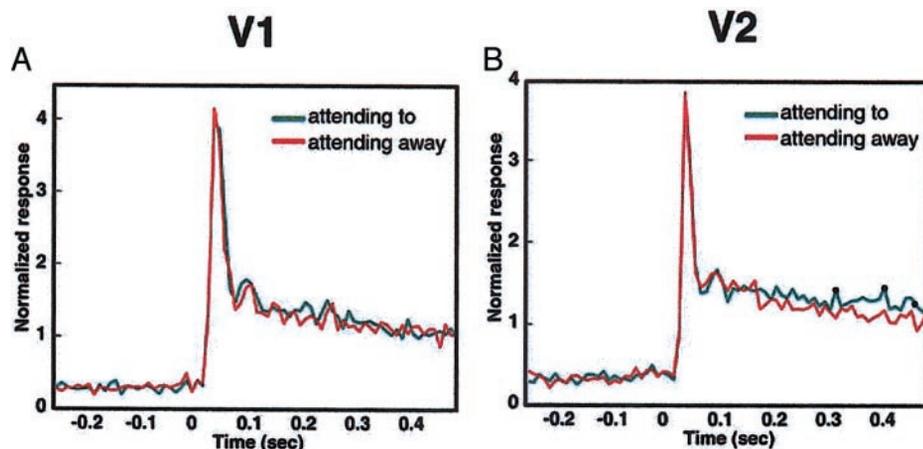


FIG. 8. Time course of attentional effects. Data was binned as described in Figs. 3–9. *A*: V1. No effect of attention was apparent at any point in the response. No bins were significantly affected by attentional condition (ranksum test, P value = 0.05). *B*: V2. Effects of attentional condition were apparent only in the later part of the response. This may be due to the monkey paying closer attention as the behavioral depth change became more imminent.

et al. 2000), assignment of border ownership (Zhou et al. 2000), and signaling of illusory contours (Peterhans and von der Heydt 1989; von der Heydt and Peterhans 1989). Our finding that enhancement in V2 occurs for parallel disk stimuli, in which segmentation cues are restricted to the illusory contour at the figural boundary, implies that this boundary information can propagate to the figure interior. The longer latency of enhancement to the parallel disk than to the orthogonal disk (approximately 150 vs. approximately 60 ms) may reflect delays related to illusory contour representation (Lee and Nguyen 2001), border ownership signaling (Zhou et al. 2000), as well as propagation of the border signal into the figural region.

Also, the enhancement was observed on the figural side of the illusory contour border but not the background side of the border, indicating that it does not reflect the proximity of a border per se but instead provides evidence that the receptive field falls within surface bounded by the illusory contour. Surfaces have been proposed as a fundamental intermediate representation in form processing (Nakayama et al. 1995). Bakin et al.'s (2000) report that many V2 neurons signal perceived surface depth rather than absolute disparity provides further evidence for V2's involvement in surface representation. This nascent surface representation may be critical for drawing attentional resources (Rensink 2000a) and for subsequent shape analysis in V4 and inferotemporal cortex. An unresolved issue is whether the subset of V2 neurons involved in surface processing are the same as or overlap with those involved in boundary processing.

Interactions between V1 and V2

In the present study, the onset of figural enhancement in V1 to the orthogonal disk had a substantially shorter latency (approximately 60 ms) than that reported by Lamme (1995) (80–100 ms). This difference in time course of enhancement may be partially accounted for by the smaller stimulus sizes used in this study (approximately 2.5 vs. 3.6° across) (Lamme et al. 1999). Another factor possibly contributing to the shorter latencies observed here is that our stimuli may have been more salient (Super et al. 2001).

Zipser et al. (1996) proposed that the long latency of contextual modulation that they observed in V1 reflects temporal delays associated with feedback from V2 and other higher areas, whereas others (Gilbert et al. 1996; Li 1999) have

suggested that horizontal connections within V1 may be the primary source of contextual modulation. The shorter latency of enhancement that we observed does not preclude a role for feedback from V2. Feedback connections from V2 have been reported to act early on the responses in V1 (Hupe et al. 2001) and to have rapid conduction velocities (Girard et al. 2001), whereas lateral connections within V1 were reported to have slow conduction velocities (Girard et al. 2001). Thus the temporal delays in modulating V1 responses may be similar for lateral projections within V1 and for feedback projections from V2. Indeed, a heterogeneous set of connections may be involved in V1 contextual effects, with feedback projections acting specifically on horizontal connections (Neumann and Sepp 1999). Such feedback may reinforce inferences made over the course of processing (Neumann and Sepp 1999; Rao and Ballard 1999). Super et al.'s (2001) report that figural enhancement is absent when figures are not detected supports this assertion.

Parallel aspects of scene segmentation

We found that figural enhancement is robust when attention is directed away from the figure over the CRF and when multiple figures are presented. This indicates that the processes associated with figural enhancement occur preattentively and in parallel across the image. It is important to reconcile these observations with psychophysical evidence that attention can play an important role in segmentation (see Driver et al. 2001). The modest effect of attention on processing of the relatively sparse and unambiguous displays used in this study suggests that the role of attention may be more prominent when segmentation requires intense scrutiny. These findings support a model of visual processing in which the scene is first segmented into discrete subregions to which attentional resources can be directed in turn (Rensink 2000a,b). When attention is directed to these “proto-objects,” a more detailed representation may then give rise to a full figure-ground percept (Scholl 2001).

General effects of attention

We observed a modest general facilitatory effect of attention in V2. The magnitude of the effect, measured over the entire response was similar in both monkeys to that previously reported in V2 (Luck et al. 1997; Motter 1993; Reynolds et al.

1999). However, in one of the monkeys, the effect was substantially stronger in the later part of the response. The longer latency may be attributable to the cueing method used in our task: instead of cueing the monkeys to a specific location as in other studies (Luck et al. 1997; Motter 1993), the monkeys were cued only to the hemifield in which the target would appear. Attentional effects may therefore have been delayed until the monkey located the figure in the cued hemifield and directed attention to it. Also, in previous studies, neuronal responses were recorded as the monkeys discriminated a cued target, whereas we recorded activity to a stimulus while the monkeys awaited the stimulus to be discriminated. The monkey may therefore have "paid more attention" as the stimulus period progressed and the discrimination stimulus became imminent. It is unclear why this late onset effect was observed in only one of the monkeys, but it may be due to the two animals adopting different behavioral strategies. This result serves as an important reminder that even well controlled behavioral tasks may have multiple solutions that are reflected differently in the neuronal responses of individual subjects.

The lack of modulatory effects of attention we observed in V1 is consistent with a number of other studies using a range of tasks (Luck et al. 1997; McAdams and Maunsell 1999). In contrast, Ito and Gilbert (1999) reported significant attentional effects when collinear flankers were presented adjacent to the CRF and suggested that attention in V1 acts by "gating the horizontal connections by feedback connections from higher order areas." On the other hand, the "pop out" stimuli by Press (1998), as well as those used here, elicited significant contextual modulation but did not yield an interaction between contextual effects and attention. Attentional effects related to stimulus context may thus be limited to a restricted range of configurations that emphasize high-resolution spatial analysis.

We thank A. Anzai, G. DeAngelis, and L. Snyder for helpful comments and discussion and H. Drury, B. Press, and E. Connor for computer programming assistance.

This research was supported by National Eye Institute Grant EY-02091.

REFERENCES

- ALLMAN J, MIEZIN F, AND MCGUINNESS E. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu Rev Neurosci* 8: 407–430, 1985.
- BAKIN JS, NAKAYAMA K, AND GILBERT CD. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J Neurosci* 20: 8188–8198, 2000.
- CHELAZZI L, MILLER EK, AND DUNCAN J. A neural basis for visual search in inferior temporal cortex. *Nature* 363: 347–347, 1993.
- CONNOR CE, PREDDIE DC, GALLANT JL, AND VAN ESSEN DC. Spatial attention effects in macaque area V4. *J Neurosci* 17: 3201–3214, 1997.
- DRIVER J, DAVIS G, RUSSELL C, TURATTO M, AND FREEMAN E. Segmentation, attention, and phenomenal visual objects. *Cognition* 80: 61–95, 2001.
- GILBERT CD, DAS A, ITO M, KAPADIA M, AND WESTHEIMER G. Spatial integration and cortical dynamics. *Proc Natl Acad Sci USA* 93: 615–622, 1996.
- GIRARD P, HUPE JM, AND BULLIER J. Feedforward and feedback connections between areas v1 and v2 of the monkey have similar rapid conduction velocities. *J Neurophysiol* 85: 1328–1331, 2001.
- HAENNY PE AND SCHILLER PH. State dependent activity in monkey visual cortex. I. Single cell activity in V1 and V4 on visual tasks. *Exp Brain Res* 69: 225–244, 1998.
- HUPE JM, JAMES AC, GIRARD P, LOMBER SG, PAYNE BR, AND BULLIER J. Feedback connections act on the early part of the responses in monkey visual cortex. *J Neurophysiol* 85: 134–145, 2001.
- ITO M AND GILBERT CD. Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* 22: 593–604, 1999.
- KAPADIA MK, ITO M, GILBERT CD, AND WESTHEIMER G. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15: 843–856, 1995.
- KNIERIM JJ AND VAN ESSEN DC. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol* 67: 961–980, 1992.
- LAMME VA. The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15: 1605–1615, 1995.
- LAMME VA, RODRIGUEZ-RODRIGUEZ V, AND SPEKREIJE H. Separate processing dynamics for texture elements, boundaries, and surfaces in primary visual cortex of the macaque monkey. *Cereb Cortex* 9: 406–413, 1999.
- LAMME VA, ZIPSER K, AND SPEKREIJE H. Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proc Natl Acad Sci USA* 95: 3263–3268, 1998.
- LEE TS, MUMFORD D, ROMERO R, AND LAMME VA. The role of the primary visual cortex in higher level vision. *Vision Res* 38: 2429–2454, 1998.
- LEE TS AND NGUYEN M. Dynamics of subjective contour formation in the early visual cortex. *Proc Natl Acad Sci USA* 98: 1907–1911, 2001.
- LI Z. Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network* 10: 187–212, 1999.
- LUCK SJ, CHELAZZI L, HILLYARD SA, AND DESIMONE R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 77: 24–42, 1997.
- MCADAMS CJ AND MAUNSELL JH. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19: 431–441, 1999.
- MOTTER BC. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol* 70: 909–919, 1993.
- NAKAYAMA K, HE ZJ, AND SHIMOJO S. Visual surface representation: a critical link between lower-level and higher-level vision. In: *Visual Cognition* (2nd ed.), edited by Kosslyn SM and Osherson DN. Cambridge, MA: MIT Press, 1995, p. 1–70.
- NELSON JI AND FROST BJ. Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Res* 139: 359–365, 1978.
- NEUMANN H AND SEPP W. Recurrent V1–V2 interaction in early visual boundary processing. *Biol Cybern* 81: 425–444, 1999.
- PETERHANS E AND VON DER HEYDT R. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *J Neurosci* 9: 1749–1763, 1989.
- PRESS WA. Effects of spatial attention on macaque primary visual cortex. In: *Computation and Neural Systems*. Pasadena, CA: California Institute of Technology, 1998, p. 170.
- RAO RP AND BALLARD DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2: 79–87, 1999.
- RENSINK RA. The dynamic representation of scenes. *Visual Cognition* 7: 17–42, 2000a.
- RENSINK RA. Seeing, sensing, and scrutinizing. *Vision Res* 40: 1469–1487, 2000b.
- REYNOLDS JH, CHELAZZI L, AND DESIMONE R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci* 19: 1736–1753, 1999.
- ROBINSON DA. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 10: 137–145, 1963.
- ROELFSEMA PR, LAMME VA, AND SPEKREIJE H. Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395: 376–381, 1998.
- ROSSI AF, DESIMONE R, AND UNGERLEIDER LG. Contextual modulation in primary visual cortex of macaques. *J Neurosci* 21: 1698–1709, 2001.
- SCHOLL BJ. Objects and attention: the state of the art. *Cognition* 80: 1–46, 2001.
- SUPER H, SPEKREIJE H, AND LAMME VA. Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat Neurosci* 4: 304–310, 2001.
- VON DER HEYDT R AND PETERHANS E. Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *J Neurosci* 9: 1731–1748, 1989.
- VON DER HEYDT R, ZHOU H, AND FRIEDMAN HS. Representation of stereoscopic edges in monkey visual cortex. *Vision Res* 40: 1955–1967, 2000.
- ZHOU H, FRIEDMAN HS, AND VON DER HEYDT R. Coding of border ownership in monkey visual cortex. *J Neurosci* 20: 6594–6611, 2000.
- ZIPSER K, LAMME VA, AND SCHILLER PH. Contextual modulation in primary visual cortex. *J Neurosci* 16: 7376–7389, 1996.