Spatial Attention Effects in Macaque Area V4

Charles E. Connor, Dean C. Preddie, Jack L. Gallant, and David C. Van Essen

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

Focal visual attention typically produces enhanced perceptual processing at the psychological level and relatively stronger neural responses at the physiological level. A longstanding mechanistic question is whether these attentional effects pertain specifically to the attended (target) object or to the region of space it occupies. We show here that attentional response enhancement in macaque area V4 extends to behaviorally irrelevant objects in the vicinity of the target object, indicating that focal attention has a strong spatial component at the physiological level. In addition, we find that spatial attention effects typically show a striking directional asymmetry. The direction of the asymmetry varies between cells, so that some cells respond best when attention is directed to the left of the stimulus, some when attention is directed to the right, etc. Thus, attention involves not only enhanced responses to behavioral targets but also a complex modulation of responses to other stimuli in the surrounding visual space.

Key words: attention; visual; spatial; monkey; cortex; extrastriate; area V4

Extensive psychological research has shown that directed visual attention produces localized enhancement of perceptual processing. In some studies, this perceptual enhancement appears to be specific to the behavioral target object, so that other objects nearby or even overlapping in space are processed less effectively (Duncan, 1984). In other studies, attentional enhancement appears to be a spatial phenomenon that applies to all stimuli in the general vicinity of the target, with the effects falling off gradually at greater distances (Posner et al., 1980). This distinction between object-based and spatial mechanisms of attention has yet to be fully addressed at the neurophysiological level. Single-unit recordings in monkeys have revealed that behaviorally relevant target objects typically evoke stronger responses than nontarget objects in extrastriate visual areas concerned with object recognition, including V2, V4, and inferotemporal cortex (IT) (Moran and Desimone, 1985; Motter, 1993). These results are generally compatible with both types of attentional mechanisms: an enhanced response to a behavioral target may represent either an object-based effect restricted to the target or a spatial effect centered on the target.

We sought to dissociate the two factors and study spatial attention effects in isolation by means of a double-stimulus paradigm. One stimulus was a ring-shaped behavioral target used to control the locus of attention but was designed and positioned so as not to evoke a response from the cell under study. The other stimulus was a behaviorally irrelevant bar designed to evoke a strong response from the cell and was used to map the spatial profile of attention effects surrounding the ring target. Thus, the spatial position of the probe stimulus was dissociated from the position of the behavioral target. The locations of both stimuli were varied systematically so as to map the spatial profile of attentional modulation. Our results reveal a penumbra of attentional enhancement surrounding the behavioral target that affects responses to nearby behaviorally irrelevant stimuli. This supports the existence of a spatial mechanism of attention (without excluding the possibility of a coexisting object-based mechanism). In addition, for most V4 cells response strength depends on the direction in which attention lies relative to the receptive field, with the optimal direction varying between cells. This indicates that area V4 carries information about the positional relationship between stimulus and attention and suggests the emergence of an attention-centered reference frame for visual processing.

A preliminary report of some of these results has been presented previously (Connor et al., 1996).

MATERIALS AND METHODS

General. All surgical, training, and neurophysiological recording procedures conformed to National Institutes of Health and USDA guidelines and were carried out under an institutionally approved animal protocol using methods described previously (Knierim and Van Essen, 1992), except as noted below.

Equipment. Visual stimuli were generated on an Iris Indigo workstation (SGI) and displayed on a 36×27 cm color monitor with a resolution of 1280×1024 pixels viewed at a distance of 55 cm. Single-unit activity in area V4 was recorded with 125-μm-diameter epoxy-coated tungsten electrodes (A-M Systems) with impedances of 1–5 MΩ. Electrodes were inserted transdurally through a 5-mm-diameter craniotomy by means of a custom guide tube system. Electrode position was controlled by a stepping motor microdrive (Caltech Central Engineering Services). The electrical waveform was thresholded with a window discriminator, and the resulting digital signal was collected through the audio input channel of the Indigo workstation at a sample rate of 8 kHz. Eye position was monitored with a scleral search coil system (Remmel Labs) using the technique of Robinson (1963). The analog eye position signals were digitized on a custom interface (Caltech Biology Electronics Shop) connected to a serial port on the workstation. This interface also processed digital input and output signals for the behavioral response lever and the liquid reward system.

Receptive field plotting. Macaque monkeys were seated before the display monitor in a standard primate chair with a response lever at-
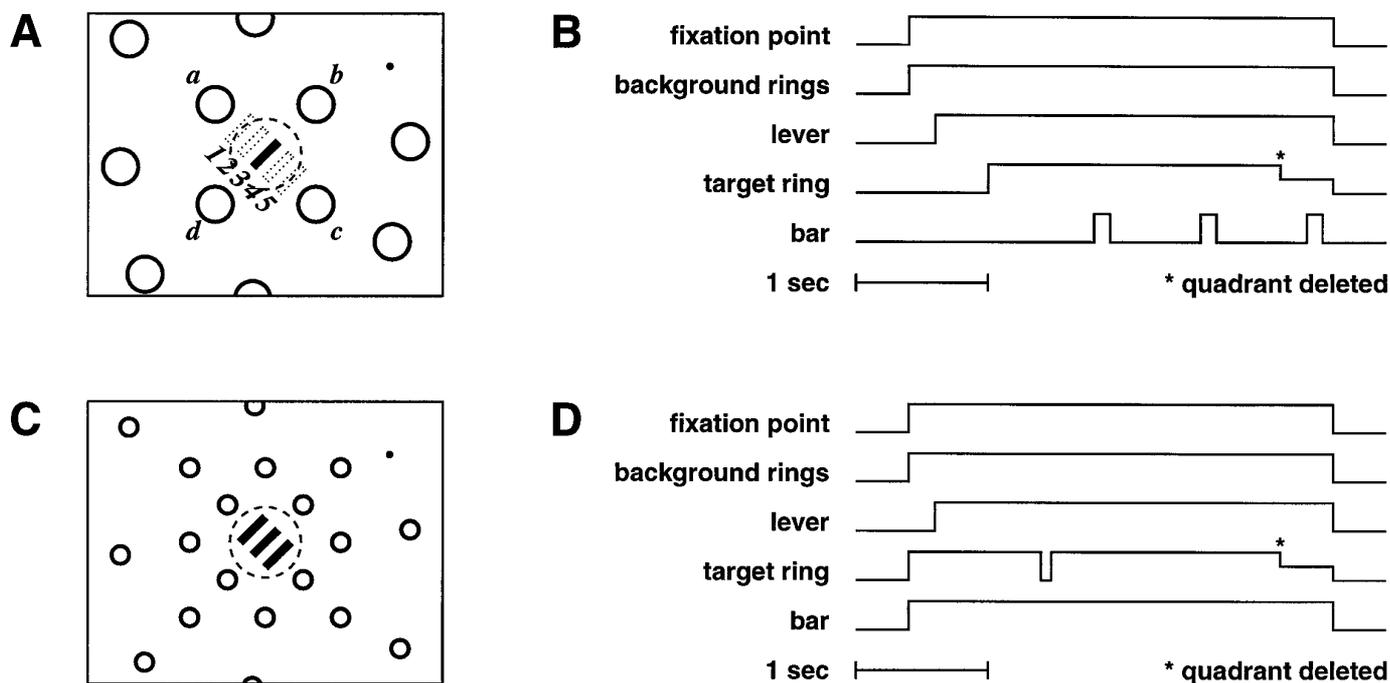
Figure 1. Stimuli and behavioral paradigms. *A,* Stimuli for the **4 ring/5 bar** test. The *dot* in the *top right* represents the fixation point, and the *dashed circle* represents the CRF. Four rings surround the CRF at positions *a–d*, and the rest of the screen is filled with distractor rings. Ring diameter is 0.5 CRF diameter. Bars are presented (individually) at positions *1–5*, which span the CRF along the axis perpendicular to bar orientation. Bar length is 0.5 CRF diameter, and bar spacing is 0.25 CRF diameter. *B,* Sequence of trial events for the 4 ring/5 bar test (example). The trial begins with the appearance of the fixation point and background rings (including rings at three of the positions near the CRF). After the animal initiates fixation and depresses the response lever, there is a 500 msec delay, and then the target ring appears at the remaining position near the CRF. The delayed onset denotes the target position for that trial. Bars are flashed individually and in random order at the five locations spanning the CRF for 150 msec each at 1 sec intervals beginning 1 sec after target onset. At a random time point no less than 500 msec after target onset, a quadrant is deleted from the target ring. The animal must respond to this by releasing the response lever. *C,* Stimuli for **12 ring** test. The CRF is surrounded by 12 potential target rings, each with a diameter equal to 0.25 CRF diameter, arranged in a square array with a spacing equal to 0.75 CRF diameter. The stimulus is a set of three bars presented simultaneously in the center of the CRF. Bar length is 0.5 CRF diameter, and bar spacing is 0.25 CRF diameter. *D,* Sequence of trial events for the **sustained bar** test. The fixation point, background rings, and target rings appear simultaneously at the beginning of the trial. The animal must initiate fixation and depress the response lever. Then, after a 1 sec delay, the target ring blinks off for 100 msec. At a random time point no less than 500 msec after the blink, a quadrant is deleted from the target ring and the animal must respond by releasing the response lever.

tached. Study of an isolated cell began with receptive field mapping. On each mapping trial, the animal was required to maintain fixation within a 0.5° diameter window for a period of 5 sec in order to receive a liquid reward. While the animal was fixating, the investigator probed the cell's responses with moving and flashing bar stimuli whose position, length, width, color, and orientation were under mouse control. Optimal values for bar width, color, and orientation were estimated. The classical receptive field (CRF) of the cell was delineated as a circular region in which reliable excitatory responses could be evoked by either drifting or flashing optimal bars. For 44 cells, this handplotting procedure was followed by an automated plotting routine in which responses were probed with optimal bar stimuli of length 0.25 CRF diameters presented in random order at locations in a square grid with a spacing of 0.125 CRF diameters and covering an area twice the CRF diameter. The bars were flashed for 150 msec each at intervals of 750 msec. The resulting data were used to find the center of gravity for response strength, and the position and size of the CRF were adjusted when appropriate.

4 Ring/5 bar test. Four different tests were used in this study. The stimulus display for the main (**4 ring/5 bar**) test is depicted in Figure 1*A*. Each trial began with the onset of a 0.1° fixation point (*dot* at *top right*) and a field of ring stimuli. The ring stimuli were sized and positioned with respect to the previously plotted CRF (*dashed circle*). The four target rings (labeled *a–d*) were half the size of the CRF and located 1.0 CRF diameter from the CRF center along the axes parallel and perpendicular to the optimal bar orientation. Distractor rings were the same size and covered the remainder of the display screen. Ring color was chosen to be nonoptimal for the cell under study, and ring line width was 0.1°. After the onset of the display, the monkey was required to initiate fixation within a 0.5° diameter window and depress a response lever. The onset of the target ring for that trial (which could be at any one of positions *a–d*)

was delayed until 500 msec after both fixation and lever press had occurred (see timeline in Fig. 1*B*). This slightly delayed onset was the behavioral cue designating the target ring to the animal. In addition, target position was blocked in sets of 12 trials so as to enhance the animal's certainty about where to attend. After target ring onset, stimulus displays were identical across target positions. The animal was required to monitor the target ring continuously for the deletion of a 90° section anywhere along the ring's circumference. This occurred at random time points between 0.5 and 4.5 sec after target ring onset. The animal received a liquid reward if it released the response lever within 700 msec of the quadrant deletion.

One-quarter of the trials were catch trials in which one or more of the other rings in the display (including nontarget rings at positions *a–d*) underwent a quadrant deletion before the target ring. The animal had to ignore these events and wait for the target ring deletion before responding. Successful performance in the catch trials demonstrated that the animal was in some sense concentrating on the target ring. The average response error rate (i.e., early or late response lever release) across the two animals was 1.4% for normal trials and 1.6% for catch trials. The average fixation error rate was 17.5% for normal trials and 17.4% for catch trials. To test whether the position of attention produced any systematic bias in eye position, average eye position associated with each of the four target ring positions was calculated for each cell and the maximum difference in eye position between any two positions was determined. The average maximum difference across all cells studied with the **4 ring/5 bar** test in the two monkeys was 0.08°. Given the range of receptive field sizes in area V4 (see below), this value is far too small to explain any of the observed response changes.

While the animal was waiting for a deletion in the target ring, bar stimuli were flashed one at a time in pseudorandom order at five locations

(labeled *1–5* in Fig. 1*A*) spanning the CRF at 0.25-CRF-diameter intervals. Bar length was 0.5 CRF diameters, and bar width, orientation, and color were set at values estimated to be optimal during receptive field plotting. The first bar flash occurred 1 sec after target ring onset, and subsequent flashes occurred at 1 sec intervals thereafter (Fig. 1*B*). Each bar was presented for 150 msec. The data analysis excluded bars that occurred within 500 msec of target ring deletion (as well as bars presented during catch trials). This meant that for a given trial anywhere between zero and three bar presentations were suitable for analysis, depending on trial length. Bar flashes continued at 1 sec intervals even after the quadrant deletion so as not to provide a secondary response cue. The different bar positions were tested equally often at the first (1 sec) time point in the trial. The same set of within-trial bar position sequences was used for each of the four target positions so as to balance any short-term order of presentation effects. The order in which these trials were presented, however, was randomized separately for the four target positions. The randomized sequences of bar positions included a condition in which no bar was presented. This provided a measure of background responses under equivalent circumstances for comparison with bar responses. In this and the other three tests, each combination of bar position and target ring position was presented eight times (in a few cases 6 times).

2 Ring/7 bar test. This procedure was used to characterize response profiles over a larger range of bar positions. The stimulus geometry was similar to that in Figure 1*A*, but only the two target ring locations on the axis perpendicular to bar orientation were tested and an extra bar position was added on either side of the CRF, with the spacing still at 0.25 CRF diameters (for a total span of 1.5 CRF diameters). The trial event sequence was similar to that in Figure 1*B*, except that the first bar presentation occurred 200 msec after target ring onset, with subsequent presentations at 1 sec intervals thereafter. Responses to the first bar presentation were not included in the analysis presented below; only bar presentations at the 1.2 and 2.2 sec time points were analyzed. Target ring quadrant deletion occurred at random time points between 0.7 and 3.7 sec after target ring onset.

12 Ring test. This procedure was used to sample attention position more densely. The sequence of trial events in this case was nearly the same as in the **4 ring/5 bar** test (Fig. 1*B*), but the stimulus geometry was different (Fig. 1*C*). The 12 potential target rings were positioned around the CRF in a square array with a spacing of 0.75 CRF diameters. Ring size was 0.25 CRF diameters. The probe stimulus was a set of three bars flashed simultaneously and spanning the central 50% of the CRF. The use of this single probe stimulus made it feasible to test the larger number of target ring positions. The three bar probe stimulus was flashed once every 2 sec beginning either 1 or 2 sec after target ring onset. The alternating 1 sec time points in which no bar stimulus was presented provided a measure of background responses. Quadrant deletion occurred anywhere between 0.5 and 4.5 sec after target ring onset. This produced zero to two analyzable bar presentations per trial. Target position was blocked in sets of four trials.

Sustained bar test. This procedure was used to study attention effects on responses to prolonged stimulus presentations. The stimulus geometry matched that in the **4 ring/5 bar** test (Fig. 1*A*), but the sequence of trial events differed (Fig. 1*D*). Each trial began with the onset of all of the rings and a single bar at one of the five positions in the CRF. One second after fixation and lever press, the position of the target ring for that trial (1 of the 4 rings surrounding the CRF) was cued by a 100 msec blink. This made it possible to examine the change in the tonic response to the continuously present bar stimulus after the appearance of a cue drawing attention to a given ring position. The animal was again required to release the response lever after a quadrant was deleted from the target ring, and to ignore deletions in other rings during catch trials. The bar stimulus remained on throughout the trial. Target ring quadrant deletion occurred at random time points between 1.0 and 3.0 sec after the blink. Target and bar position (including a no bar condition) were randomized across trials, with no blocking.

Data collection. Data were obtained from 205 area V4 cells studied with one or more of the tests described above. One-hundred and five cells were studied with the **4 ring/5 bar** test, 52 cells with the **2 ring/7 bar** test, 29 with the **12 ring** test, and 39 with the **sustained bar** test. The cells were sampled from the right hemisphere of a female *Macaca nemestrina* and the right and left hemispheres of a male *M. mulatta*. The cells were recorded from the lower visual field representation in the prelunate gyrus and adjoining banks of the lunate and superior temporal sulci. Assignment of cells to area V4 was based on the local retinotopic pattern of visual field representations, receptive field size, stimulus response prop-

erties, and location of the electrode penetrations on the skull in relation to other experiments in which the position of area V4 was verified histologically. Receptive field eccentricity ranged from 1.5° to 14°, and receptive field diameter ranged from 1.4° to 10°.

General data analysis. The response to each flashed bar presentation in the 2 and 4 ring tests, as well as the automated receptive field test, was measured by counting spikes within a 450 msec analysis window beginning at bar onset. A shorter, 250 msec window was used to analyze results in the **12 ring** test because responses in this test were substantially weaker and more transient, perhaps because of a greater inhibitory influence from the denser array of surround stimuli. These analysis windows were large enough to encompass all stimulus responses regardless of latency differences and despite a ±50 msec uncertainty in the temporal alignment between the stimulus event record and the spike event record caused by a numerical truncation error discovered after the data had been collected. The temporal uncertainty does not significantly impact the results in this paper, which depend solely on the overall response to each bar flash and not on the precise temporal relationship between stimulus events and spikes. Analysis was restricted to bar presentations beginning at least 500 msec before target ring deletion (and catch trials, with deletions at other ring positions, were excluded entirely), so that the analysis windows never overlapped stimulus changes elsewhere in the display.

Two analysis windows were used to measure responses in the **sustained bar** test. One was a 500 msec window beginning at the onset of the blink that cued target ring position. This early analysis window was designed to include any phasic response changes that might occur immediately after cueing. It has the disadvantage of possible stimulus effects attributable to the target ring blink (although this occurred outside the plotted CRF). For this reason, a late analysis window beginning 500 msec after the start of the blink (400 msec after the end of the blink) and lasting 450 msec, designed to capture tonic response changes well after the attentional cue, was also used.

In all tests, response rates were also calculated for the randomly interspersed time windows in which no bar stimulus was presented. This provided a measure of background activity under conditions otherwise equivalent to bar presentation conditions. All analyses described below were performed both with and without previous subtraction of average background rates, with similar results. Single-cell examples throughout the paper show unadjusted response rates (no background subtraction); in most cases, background rates are also indicated. Population results are based on previous background subtraction.

Shift analysis. We analyzed two kinds of response changes that occurred as the position of attention was varied. One was a shift in the response profile of the cell, typically toward the target ring. Because only the axis perpendicular to bar orientation was probed at multiple positions, the critical target ring positions for this analysis were the two positions along this axis (positions *a* and *c* in Fig. 1*A*). The shift in response profile was quantified in two ways, as follows. (1) **Shift in peak response position**. This was the change in bar position evoking the strongest response, measured in CRF diameters. The shift was considered positive if it followed the direction of attention. This measure had a potential range of −1.0 to 1.0 CRF diameters in the **4 ring/5 bar** test and −1.5 to 1.5 diameters in the **2 ring/7 bar** test. (2) **Fractional shift of response strength**. This was the proportion of total response strength that shifted from one-half of the CRF to the other as attention was directed from one side to the other. It was calculated as follows:

$$\text{Fractional shift} = \frac{R_1^a + R_2^a}{R_1^a + R_2^a + R_4^a + R_5^a} - \frac{R_1^c + R_2^c}{R_1^c + R_2^c + R_4^c + R_5^c},$$

where R_x^i is the average response rate at bar position *i* when attention is directed to ring *x*. In effect, the fractional shift index is a comparison of aggregate response strength at bar positions 1 and 2 when attention was directed to ring *a* versus ring *c*. The same value would be obtained if the equivalent comparison were made for bar positions 4 and 5. Bar position 3 was excluded from this analysis because it fell at the center of the CRF. For the **2 ring/7 bar** test, each half-total included three bar positions rather than two. Like the peak shift index, the fractional shift index yielded a positive value for response shifts toward attention. The highest possible value of 1.0 would indicate a shift of the entire response from one-half of the receptive field to the other. The lowest possible value of −1.0 would indicate an equivalent shift in the opposite direction. A value

of 0.0 would represent no change in the relative response rates in the two halves of the response profile.

Statistical significance of response profile shifts was determined with a two-tailed randomization test (Manly, 1991). The test statistic was the fractional shift index, and the null hypothesis was that the fractional shift was 0, i.e., that shifting attention from one side of the CRF to the other did not change relative response strength in the two halves of the CRF. The obtained fractional shift value was compared to a distribution generated by randomly permuting response rates across target positions *a* and *c* (within bar position) and recalculating the fractional shift 10,000 times. If the original fractional shift fell within the upper or lower 2.5% of the randomized distribution, the effect was considered significant at the 5% level.

Directional analysis. The second kind of response change requiring quantitative analysis was modulation of total response strength depending on target ring position. Total response strength for each target ring position was calculated by summing the mean response across all five bar positions (or, in the **12 ring** case, just determining the mean response to the 3 bar stimulus). The modulation of total response strength with target position was characterized by the fractional gain between the highest and lowest values:

$$\text{Fractional gain} = \frac{T_{max} - T_{min}}{T_{max}},$$

where T_{max} is the highest total response strength associated with an individual target position and T_{min} is the lowest. The lowest possible fractional gain value of 0.0 would indicate that response strength was completely unaffected by target ring position. The highest possible value of 1.0 would indicate that responses were completely absent for at least one target ring position.

Statistical significance of total response modulation was determined with a one-factor (target ring position) randomization ANOVA (Manly, 1991). The null hypothesis was that target ring position did not affect total response strength. The test statistic was calculated by finding the response total for each target position (see above), squaring, and summing the squared response totals across the four target positions (equivalent to an *F* ratio). The obtained test statistic was compared with a distribution generated by randomly permuting response values across all target positions (within bar position) and recalculating the statistic 10,000 times. If the original value fell within the upper 5% of this distribution, the modulation effect was considered significant. The same randomization procedure was used to calculate the fractional gain index expected on the basis of random variation in responses.

RESULTS

Our experiments were designed to characterize how area V4 responses depend on the spatial relationship between the stimulus driving the cell and the attentional locus. The results revealed two basic phenomena relevant to neural mechanisms of focal attention. One phenomenon was a shift in the cells' response profiles toward the attended location, reflecting a spatial window of enhanced responsiveness centered on but extending beyond the behavioral target. The other phenomenon was an overall change in response strength depending on the directional position of attention relative to the CRF, suggesting that V4 cells encode not only the retinotopic position of the stimulus but also its position relative to the center of attention. These two phenomena typically occurred together but will be discussed separately.

Response profile shifts

Figure 2 presents results for a V4 cell showing strong response profile shifts in the direction of attention. This cell was broadly tuned for bar orientation, so the attention-related shifts could be studied along two orthogonal axes: 45° clockwise (Fig. 2*A*) and 45° counterclockwise (Fig. 2*B*) from horizontal. Each histogram in Figure 2 shows the average responses to bar stimuli flashed at five locations spanning the CRF (shown as a *dashed circle* in the iconified stimulus displays). The different histograms are based on responses collected as the monkey attended to different target
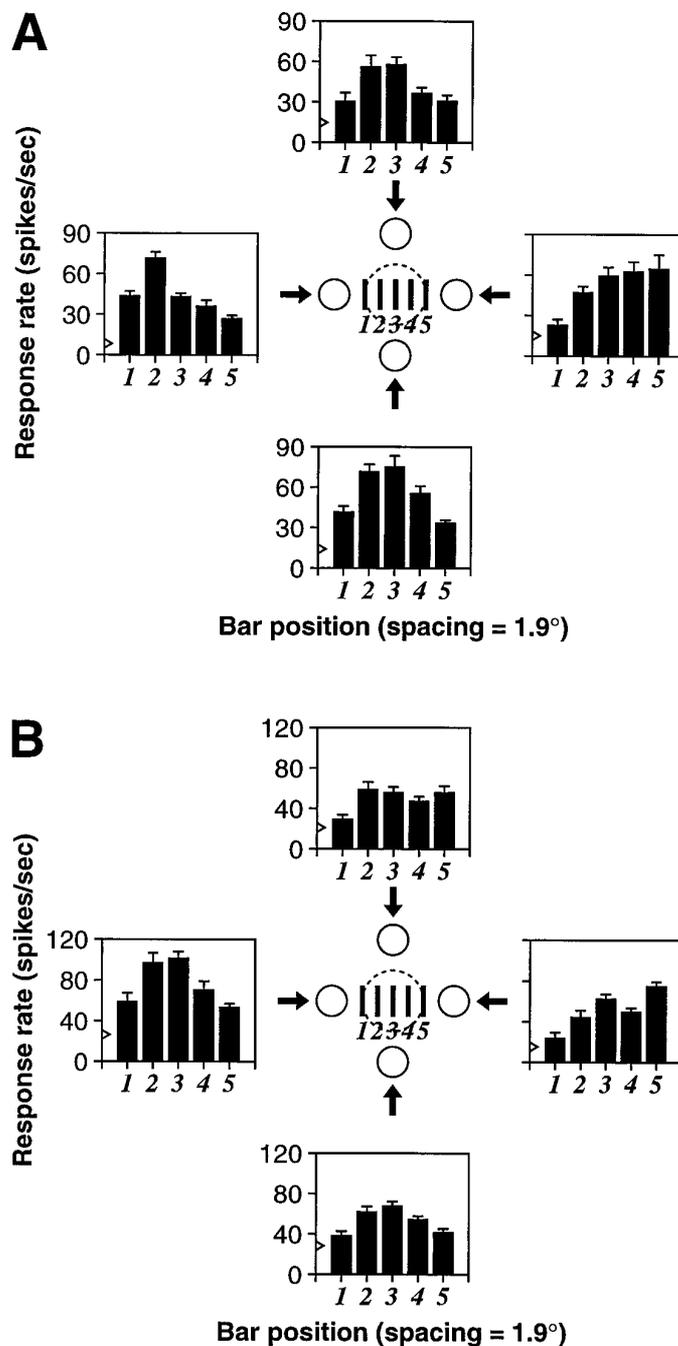

Figure 2. Response profile shift (example). This cell's CRF had a diameter of 7.5° and was located 6.7° to the left and 8.8° below fixation. It responded to red bars at all orientations, so the response shift was tested along two axes, 45° clockwise (*A*) and 45° counterclockwise (*B*) from horizontal. Ring color was green. The *center diagrams* show the bar positions and the four target ring positions. Each histogram shows the mean and SE of responses at the five bar positions when attention was directed to the target ring indicated by the *arrow*. Background response rates are indicated by the *arrowheads* along the vertical axes. In both *A* and *B*, response profiles are shifted in the direction of the target ring.

ring positions (indicated by the *arrows*) outside the CRF. Figure 2*A* shows that the region of strongest response shifted toward the target ring as attention was directed to the left or right (in reality, to the upper left or lower right). When the response profile was mapped along the orthogonal axis (lower left vs upper right, Fig.
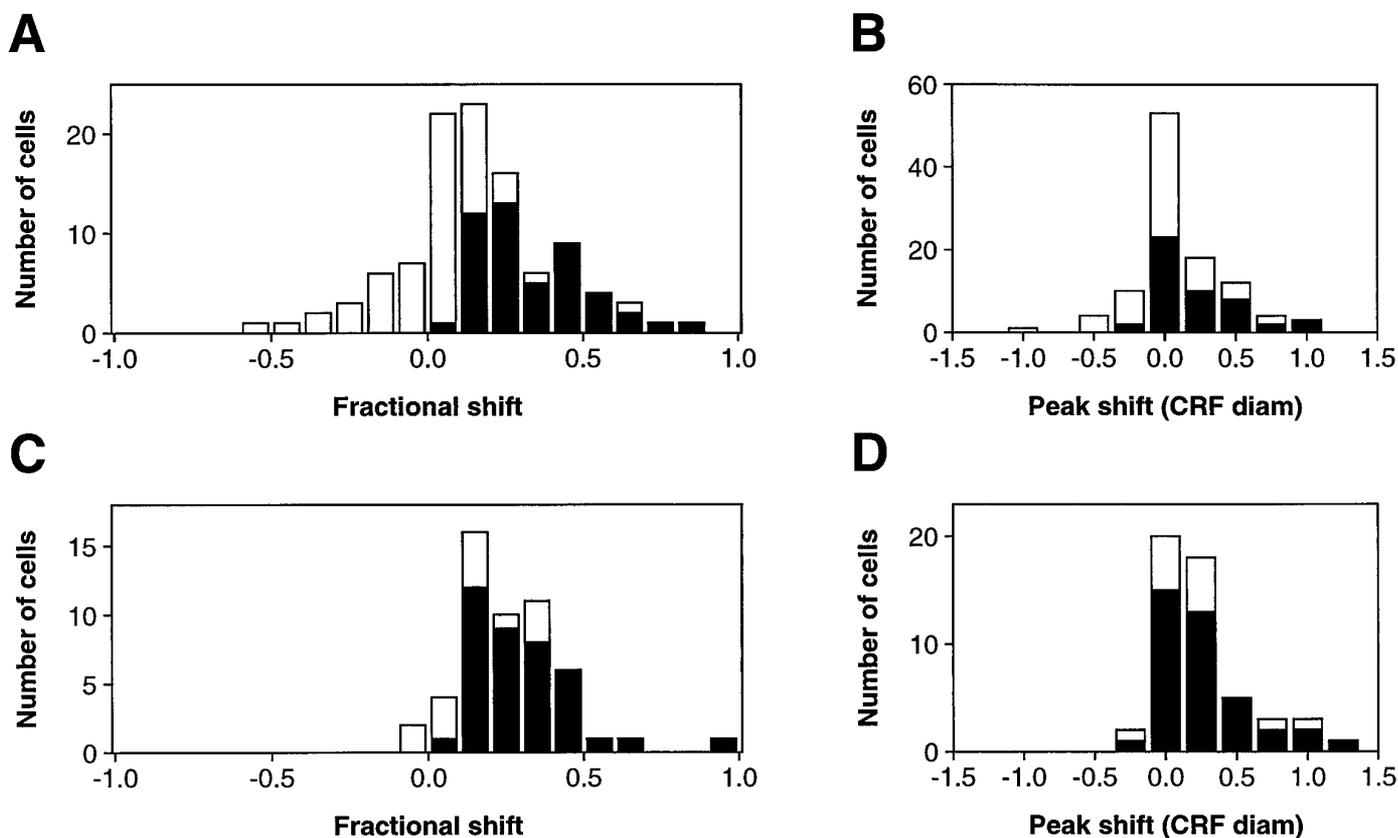
Figure 3.   Response profile shifts (population analysis). *A,* Fractional shift distribution in the **4 ring/5 bar** test. The fractional shift is the proportion of the total response profile that shifted from one side of the CRF to the other with target ring position (see Materials and Methods). Positive values represent a shift toward the target ring, and negative values represent a shift in the opposite direction. The distribution is heavily weighted in the positive direction. Values significant at the 5% level according to a randomization test are shown in *black*. None of the negative values was significant, indicating that those cells showing large negative shifts also had highly variable responses. *B,* Peak shift distribution in the **4 ring/5 bar** test. The peak shift is the distance by which the bar position producing maximum responses shifted with target ring position (see Materials and Methods). Positive values represent a shift toward the target ring, and negative values represent a shift in the opposite direction. Values associated with significant fractional shifts are shown in *black*. *C,* Fractional shift distribution in the **2 ring/7 bar** test. *D,* Peak shift distribution in the **2 ring/7 bar** test.

2*B*), a comparable response profile shift was observed. Thus, there was a consistent modulation of response strength in the space surrounding the behavioral target: Responses were relatively stronger near the target ring and weaker at greater distances.

Response profile shifts were quantified by measuring the fraction of a cell's total response strength that shifted from one side of the plotted CRF to the other under the influence of attention (see Materials and Methods). For the results presented in Figure 2*A*, this fractional shift value was 0.39; in other words, 39% of the total response strength shifted from the upper left of the CRF to the lower right. The distribution of fractional shift values for 105 cells studied with the **4 ring/5 bar** test is shown in Figure 3*A*. Positive values on the *x*-axis correspond to response shifts in the direction of attention; negative values correspond to shifts away from attention. The average fractional shift across this population was 0.16. Significance of response shifts was assessed with a two-tailed randomization test (see Materials and Methods). Forty-six percent of cells (48/105 cells) showed a significant ($p <$ 0.05) response shift in the direction of attention, and no cells showed a significant shift in the opposite direction. Significant effects are plotted in black in Figure 3*A*. The average fractional shift among cells showing significant effects was 0.33.

Response shifts were also quantified by measuring the displacement of the peak bar response position in CRF diameters. Given the stimulus geometry, the peak shift value for an individual cell was constrained to be a multiple of 0.25. The distribution of peak shifts for the 105 cells studied with the **4 ring/5 bar** test is shown in Figure 3*B*. Again, positive values represent shifts toward attention and negative values represent shifts away from attention. The average peak shift was 0.10 CRF diameters. Cells plotted in black are those that showed a significant fractional shift in the analysis presented in Figure 3*A*. Two cells showed a significant fractional shift toward attention but a -0.25 CRF diameter peak shift (away from attention). The average peak shift among cells showing significant effects was 0.22 CRF diameters.

The phenomenon of response profile shifts was studied further in 52 cells using the **2 ring/7 bar** test (see Materials and Methods), with five bar positions spanning the CRF as before and two located 0.25 diameters outside the CRF in either direction. The distribution of fractional shifts for this sample (Fig. 3*C*) tended toward larger values, perhaps because of the extended test region. The average fractional shift was 0.26, and 39 of 52 cells (75%) showed a significant ($p <$ 0.05) effect. The distribution of peak shift values is shown in Figure 3*D*; the average was 0.25 CRF diameters.

A more detailed picture of the average response profile shift is provided in Figure 4. For each cell, responses were normalized by dividing by the maximum response (i.e., the mean response asso-
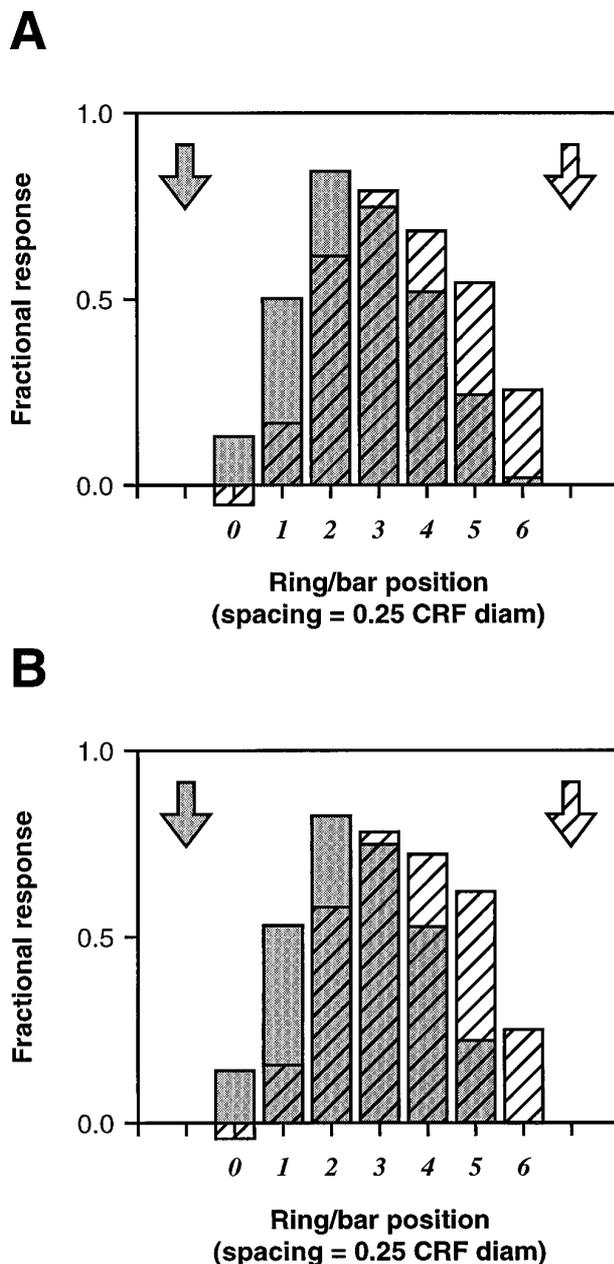
A

B

Figure 4. Average response profiles in the **2 ring/7 bar** test. For each cell, average responses at the seven bar positions were normalized by dividing by the maximum average response. Normalized responses were averaged across cells for each bar position/target ring combination. Responses associated with the two target rings are differentiated by the *halftone* and *striped patterns*. The two target ring centers are denoted by *arrows*. Average response profiles for the entire population of 52 cells are shown in *A*, and average response profiles for the 39 cells with significant fractional shift values are shown in *B*.

ciated with the most effective bar position/target ring combination). The resulting normalized values were averaged across cells. Averaging across the 52 cells tested with the **2 ring/7 bar** test produced the response profiles shown in Figure 4*A*. The arrows indicate target ring positions, and the corresponding bar responses are plotted with matching stripe and halftone patterns. The largest absolute response differences were obtained when the bar appeared at the edge of the CRF (positions *1* and *5*), whereas the greatest percentage change occurred when the bar appeared

outside the CRF (positions *0* and *6*). A similar plot based on just the 39 cells showing significant shift effects in the **2 ring/7 bar** test is shown in Figure 4*B*; the response differences are slightly larger.

Interpreting the response profile shift as an attentional phenomenon depends on discounting the possibility of stimulus-related effects. Although the stimuli in this experiment were exactly equivalent across conditions at the time bar responses were tested, the delayed onset of the target ring (for purposes of behavioral cueing) introduced an antecedent stimulus difference into each trial. It might be argued that the response profile shift is a long-term sensory effect attributable to target ring onset. If this were so, however, one would expect the effect to degrade with time, so that response shifts would be stronger at the first time point for testing bar responses (1 sec after target ring onset) and weaker at the later time points (2 and 3 sec after target ring onset). In fact, the average response profile shift was marginally stronger at the later time points (fractional shift = 0.163 in the **4 ring/5 bar** test) than at the first time point (fractional shift = 0.155). Thus, the shift effect is unlikely to be sensory in nature.

Another interpretational issue is the effect of the bar flashes themselves on the attentional state of the animal. Sudden-onset stimuli like the flashed bars used here tend to capture attention automatically (Yantis, 1993). Attentional capture by the bar flashes cannot by itself explain the response profile shifts, because the bar stimuli were completely balanced across target ring conditions. Although not sufficient to explain response shifts, attentional capture might nevertheless be a necessary factor; in other words, response profile shifts might be limited to the case of sudden-onset stimuli. Alternatively, attentional capture by the bars could have drawn attention away from the target ring and actually diminished the response profile shift. This issue was addressed with a **sustained bar** test in which the standard four target ring and five bar positions were probed, but in each trial a single bar was present throughout the trial along with a full array of rings (see Materials and Methods and Fig. 1*D*). One second into the trial, one of the rings was designated as the target with a 100 msec blink. The effect of this attention cue on the tonic response to the bar stimulus was examined. Thirty-nine cells with consistent tonic responses were studied with this procedure. Responses were analyzed within two time windows: a 500 msec window beginning at blink onset and a 450 msec window beginning 500 msec after blink onset. The 0–500 msec window was designed to encompass transient attentional effects, and the 500–950 msec window was designed to capture sustained effects. Both analysis windows yielded results comparable to the flashed bar results. The average fractional shift across the 39 cells for the 0–500 msec window was 0.29. The average for the 500–950 msec window was 0.33. In both cases, 16 of 39 (41%) of the fractional shift values were significant at the 5% level. These results argue that sudden-onset stimuli are not essential to the response profile shift effect. The larger average fractional shifts (~0.3) compared to the flashed bar average (0.16) suggest that the sudden bar onsets may indeed have diminished the effects of voluntary attention to the target rings.

Directional asymmetry

A quantitatively stronger phenomenon revealed by these experiments was directional asymmetry in attentional modulation, an example of which is shown in Figure 5. The format is similar to that in Figure 2, although here the actual bar orientation was 15° clockwise from vertical. This cell responded well to bar stimuli when attention was directed to the target ring below the CRF.
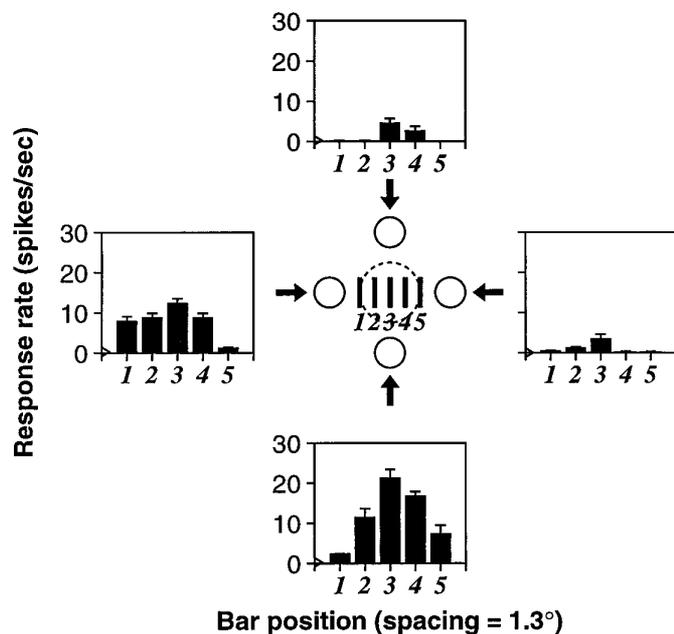
Bar position (spacing = 1.3°)

Figure 5. Directional asymmetry (example). This cell's CRF had a diameter of 5.3° and was located 4.0° to the right and 5.5° below fixation. Bar orientation was 15° clockwise from vertical. Responses were tested with magenta bars; ring color was green. As in Figure 2, the individual histograms show mean bar responses and SEs associated with the four target ring positions. Background responses are indicated by *arrowheads* along the vertical axes. Responses were strong when attention was directed to the target ring below the CRF, moderate when attention was directed to the left, and weak when attention was directed above or to the right.

Fractional gain

Figure 6. Fractional gain distribution in the **4 ring/5 bar** test. Fractional gain is a measure of the directional asymmetry in spatial attention effects. It corresponds to the proportional difference in total response strength (summed across bar positions) between the two target positions producing the maximum and minimum responses (see Materials and Methods). Values for cells that showed effects significant at the 5% level according to a randomization ANOVA are shown in *black*. The *arrow* indicates the average fractional gain expected on the basis of random variation.

Responses were weaker when attention was directed to the left and were almost absent when attention was directed above or to the right. Thus, this cell showed strong tuning for the position of attention relative to the CRF. A modest response profile shift is also apparent.

Directional asymmetry was quantified by summing responses across bar positions for each target ring condition and calculating the fractional gain between the largest and smallest summed responses (see Materials and Methods). For the example cell in Figure 5, this fractional gain was 0.90; in other words, there was a 90% drop in total response when attention was directed to the right as opposed to below the CRF. The distribution of fractional gain values for 105 cells studied with the **4 ring/5 bar** test is shown in Figure 6. A value of 0 on the *x*-axis corresponds to no directional effect, a value of 0.5 corresponds to a twofold difference between target conditions, and a value of 1.0 corresponds to a complete absence of responses in one target condition. The average expected value for fractional gain, based on the measured variability of responses across repetitions, was 0.26 (*arrow* in Fig. 6). The obtained values were largely distributed above this expected average; the obtained average was 0.55. The significance of the variation in response strength with target ring position was assessed with randomization ANOVA tests (see Materials and Methods). Eighty-five percent of cells (89/105 cells) showed effects significant at the 5% level. Cells showing significance in this test are plotted in black in Figure 6.

The foregoing analyses are independent of the actual spatial positions of the target rings, so significant results do not necessarily imply a single preferred direction. An alternative would be a multi-lobed effect (for example, strong responses when attention
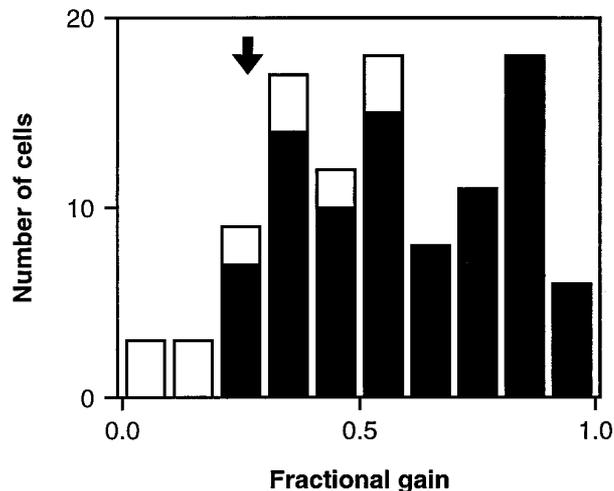
is directed to the right or left and weak responses when attention is directed above or below the CRF). This was not observed to any great extent, however. The target ring positions producing the two largest responses were nonadjacent for only 21 of 105 cells, and these were mostly cases in which the second and third largest responses were only marginally different (difference < 20% of the maximum response in all but 4 cases). Thus, the predominant pattern was one in which increased response strength was associated with a single direction.

Directionality was assessed for each cell by calculating a vector sum of the four target ring directions weighted by the normalized response strength associated with the corresponding target (see legend to Fig. 7). The magnitude of the vector sum, which varies between 0 and 1.0, reflects the degree of directional asymmetry. The vector angle estimates the optimum attention direction for the cell. Figure 7 presents the results of this analysis. Each cell is represented by an arrow, the position of which with respect to the fixation point (the intersection of the degree scales) corresponds to the CRF center and the angle and size (area) of which correspond to the angle and magnitude of the vector sum. Cells with vector magnitudes significant at the 5% level according to a randomization test are plotted in black; this included 89% of the sample (93/105), slightly higher than the percentage obtained with the (nondirectional) ANOVA (85%). The example cell shown in Figure 5 is represented by the large arrow located just under the asterisk and pointing to the lower left. The vector magnitude for this cell was 0.55.

The overall distribution of vector angles in Figure 7 appears approximately random, but two specific hypotheses regarding directional bias were considered. One hypothesis was that the direction of asymmetry might be biased relative to the horizontal or vertical meridian. This possibility is explored graphically in the polar plot of Figure 8*A*, where the angle of each dot represents vector direction from Figure 7 and distance from the center represents vector magnitude. In this plot, right/left has been transformed into ipsilateral/contralateral, because this seemed a
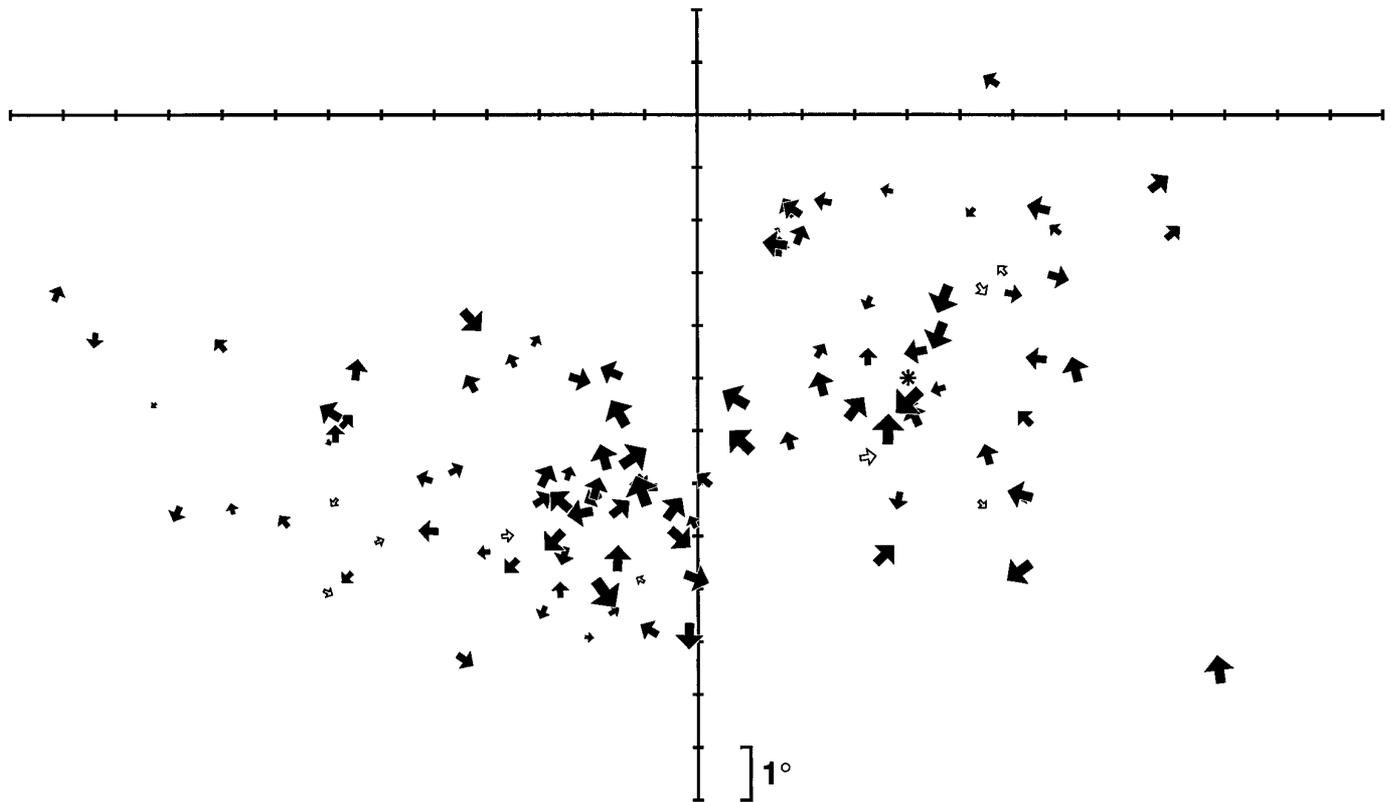
Figure 7. Vector sums in the **4 ring/5 bar** test. For each cell, an *arrow* is used to represent the vector sum of the normalized response strengths associated with the four target positions. Each component vector in the sum originates at the CRF center, points toward the appropriate target ring, and has a magnitude proportional to the total response strength when attention was focused on that target ring. The sum of these vectors reflects the directionality of the attentional profile. A cell that responded only when attention was focused on one of the rings would have a large summed vector pointing in that direction, whereas a cell that gave equivalent responses under all attention conditions would have a summed vector of negligible magnitude. The magnitude of the summed vector was normalized by the sum of the magnitudes of the component vectors (i.e., by the cell's total response strength), so that the final value would fall between 0.0 and 1.0. This normalized magnitude is represented in each case by the area of the *arrow* in the plot. The position of the *arrow* indicates the CRF position relative to fixation (the intersection of the degree axes), and the direction of the *arrow* indicates the direction of the vector sum. Normalized vector magnitude was used as a randomization statistic to test the null hypothesis that response strength was not related to the direction of attention relative to the CRF. A distribution was generated by randomly permuting response values across target ring position (within bar position) and recalculating the normalized vector magnitude 10,000 times. If the original magnitude fell within the top 5% of this distribution, the relationship between response strength and direction of attention was considered significant. Significant vectors are plotted in *black*. The *large arrow* pointing to the *bottom left* just *below* the *asterisk* corresponds to the example cell of Figure 5.

more likely dimension for directional bias. The scales are marked in increments of 0.1. The example cell shown in Figure 5 is again indicated by an asterisk. The spatial average over all the vectors (*circle*) deviates slightly in the superior and contralateral directions (by 0.05 and 0.03, respectively). The upward deviation is significant according to a standard t test ($p < 0.01$), but the contralateral deviation is not ($p > 0.05$).

The other hypothesis was that the direction of asymmetry might be biased with respect to the fovea. In Figure 8*B*, the vectors have been transformed such that the vertical axis represents the foveal/peripheral axis. The spatial average in this case shows a slight but significant deviation of 0.05 toward the fovea ($p < 0.01$). The deviations in the foveal and superior directions may represent a single phenomenon, because most receptive fields were in the lower visual hemifield. The foveal/superior bias does not in any sense constitute the entire effect, because there are numerous data points in all quadrants. However, it reflects a significant weighting in favor of stimuli that are more peripheral (or inferior) with respect to the attentional focus.

The directional asymmetry in attentional modulation was studied in greater spatial detail with a **12 ring** test (see Materials and Methods and Fig. 1*C*). Results from this test for the same example cell presented in Figure 5 are shown in Figure 9*A*. The responses to the 3 bar stimulus presented in the CRF (*dashed circle*) are represented by the heights of the rectangular blocks, which are ruled along their sides in increments of 2 spikes/sec. Each block represents the average response obtained as the animal attended to the target ring at that position. For the sake of visibility, the plot has been rotated so that target positions above the CRF are toward the front in the figure and target positions below the CRF are toward the back. Consistent with the results in the **4 ring/5 bar** test, the cell responded well when attention was directed below the CRF (toward the *back* in Fig. 9*A*) or to the lower left (*back right* in Fig. 9*A*). Responses were weak or absent when attention was directed elsewhere. Background responses associated with each target ring position are shown in Figure 9*B*.

A variety of modulation profiles was observed among the 29 cells studied with the **12 ring** test. Whereas some cells showed a fairly sharp tuning peak, as in Figure 9*A*, others were much more broadly tuned, as exemplified in Figure 9*C*. And whereas some cells showed a peak close to the CRF (Fig. 9*A,C*), in other cases the peak was at one of the most distant positions tested (Fig. 9*D*).
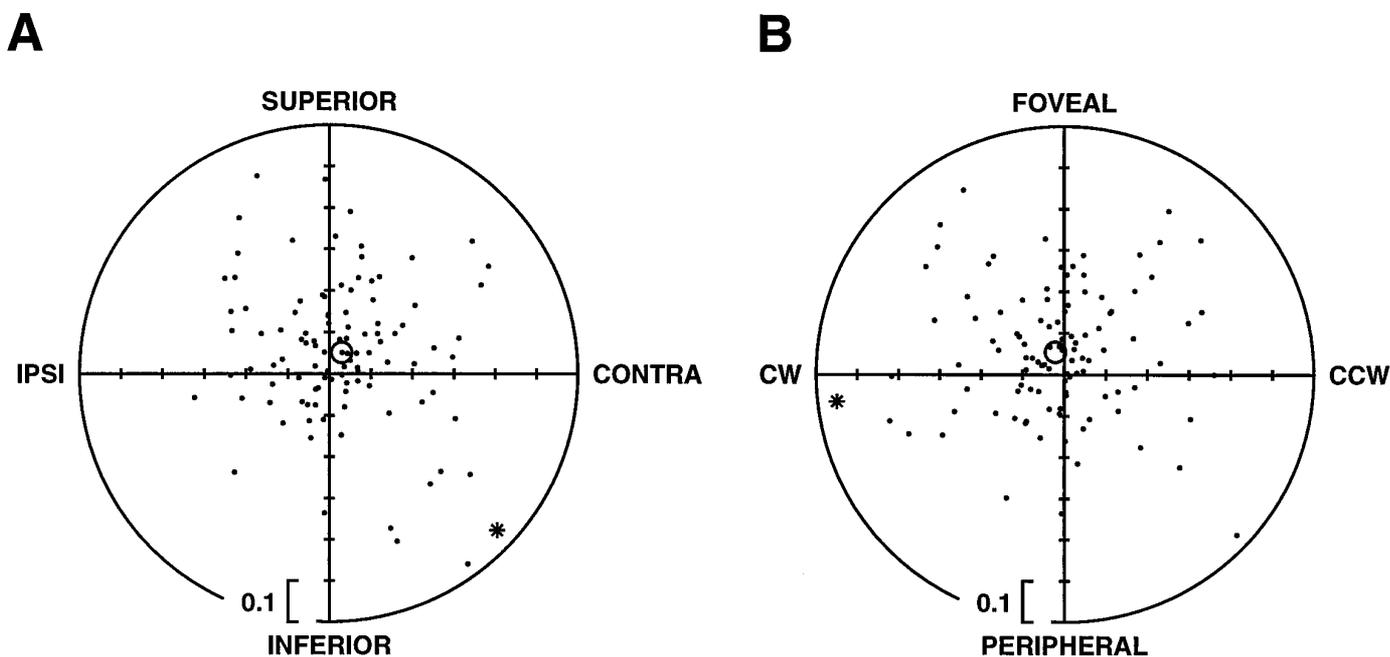
A

B

Figure 8. Distribution of vector sum directions in the **4 ring/5 bar** test. *A,* Vector direction in gravitational coordinates. Each normalized vector sum from Figure 7 is plotted as a *dot,* the angle of which represents direction in gravitational coordinates and the distance from the center of which represents magnitude. *IPSI,* Ipsilateral; *CONTRA,* contralateral. The *asterisk* corresponds to the example cell from Figure 5. The *circle* shows the spatial average across all the vectors. *B,* Vector direction in foveocentric coordinates. The angle of each *dot* represents vector direction relative to the axis running from the CRF center to the fixation point. *CW,* Clockwise; *CCW,* counterclockwise. The *asterisk* corresponds to the example cell from Figure 5. The *circle* shows the spatial average across the vectors.

Breadth of tuning was quantified by determining the fraction of target ring positions for which responses to the bar stimulus were at least half of the maximum response. This fraction was 0.17 for the cell in Figure 9A and 0.58 for the cell in Figure 9C (after background subtraction). The distribution of this breadth metric across 29 cells is shown in Figure 10A. Cells with significant ($p < 0.05$) attentional modulation according to a randomization ANOVA are plotted in black; these included 15 of 29 cells (52%). An analysis of tuning peak distance is presented in Figure 10B. This analysis was restricted to the eight target ring positions along the axes parallel and orthogonal to the stimulus bars, because these constituted a complete test of 4 directions × 2 distances (0.75 and 1.5 CRF diameters from the CRF center). The other four target positions confounded distance and direction. The distribution of peak distances was weighted toward the smaller value (Fig. 10B). Thus, responses tended to be stronger when the attended target was moved closer to the stimulus and the CRF. This is consistent with the response shift effect (see above), which showed, conversely, that responses tend to be stronger when the stimulus is moved closer to the attended target.

One obvious factor that might contribute to the directional effect is asymmetry in the strength of the surround region outside the excitatory portion of the CRF (cf. Allman et al., 1985; Desimone et al., 1985). Directing attention to stronger or weaker parts of the surround might modulate responses to stimuli inside the CRF. (This would not alter the attentional nature of the phenomenon; see Discussion.) This hypothesis was tested in 41 cells for which both an automated receptive field plot and the **4 ring/5 bar** data were available. We compared the receptive field plot responses in the regions of the four target rings with the corresponding responses in the **4 ring/5 bar** test. The correlations ranged from −0.99 to 0.77, with an average of −0.15 for all 41 cells.

(Extreme individual correlation values are not surprising with only 4 data points, especially if a few points cluster, as was often the case here.) Some examples are shown in Figure 11. Response magnitudes in the automated plot are represented by gray levels at each spatial location. The dashed ring indicates the plotted CRF, and the surrounding target rings are shown in white. At each target location, the summed bar response from the **4 ring/5 bar** test ("bar") and the average response from the receptive field plot ("RF") are given in spikes/sec. In Figure 11A–C, there are clear inhibitory zones (representing suppression of background firing in the automated receptive field plot) in the vicinities of the target rings. The correlation between this inhibition and the **4 ring/5 bar** responses ranged from near 0 (Fig. 11A) to negative (Fig. 11B) to positive (Fig. 11C). In Figure 11D, the surround responses were mainly excitatory and the correlation is again small (0.15), even though this cell had a particularly large fractional gain (0.90). In general, the correlation remained inconsistent for cells with large effects; for 28 cells tested with the automatic plotting procedure and showing fractional gains > 0.5, the average correlation was −0.20. The overall lack of consistent correlation argues against a simple relationship between the directional attention effect and receptive field surround structure.

An alternate approach to investigating the nonclassical surround mechanism might be to measure responses to the onset of the target ring (which appeared 500 msec after the beginning of the trial). Unfortunately, most cells showed a transient surge in background rates associated with the beginning of the trial, which tended to obscure any response to target onset. Of the 37 cells with no such initial background firing (<2 spikes/sec in a 300 msec window beginning 100 msec after trial start), 28 showed no appreciable modulation associated with target onset (<2 spikes/sec change in a 500 msec window beginning at target onset). Thus,
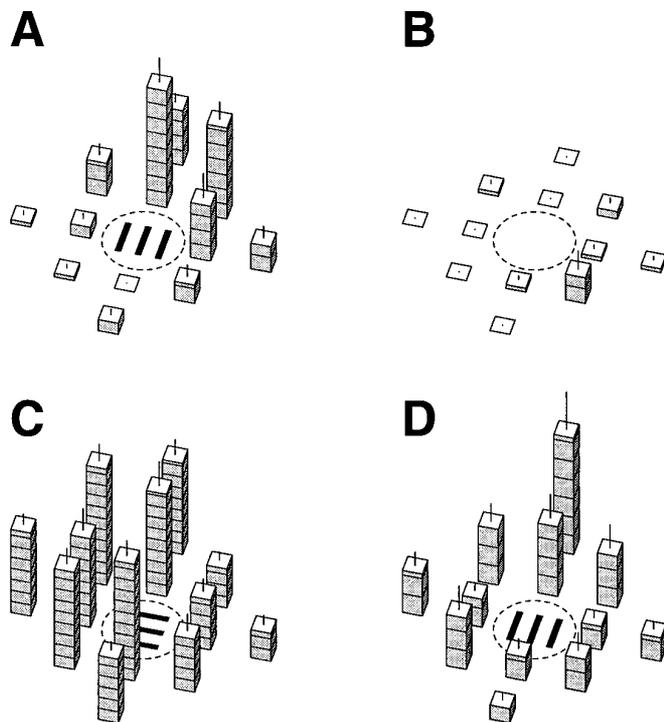
Figure 9. Directional asymmetry in the **12 ring** test (examples). *A,* Results of the **12 ring test** for the cell shown in Figure 5. The CRF is represented by the *dashed circle* with the three bar stimulus inside. The average response to the three bar stimulus as the animal attended to each of the 12 target ring positions is indicated by block height at the corresponding position. The blocks are ruled in increments of 2 spikes/sec, and the SE is indicated by the projecting *line*. The plot has been rotated so that target positions above the CRF are toward the front and target positions below the CRF are toward the back. *B,* Background responses for the same cell. *C, D,* Other examples.

there was little in the way of a measurable response to the target rings that could be used to assess the receptive field surround. Of the 28 cells with no background surge or target response, 20 had fractional gain values exceeding 0.5. This indicates that the directional effect was not generally associated with a differential response to the target ring.

The effect of target ring onset was also assessed by examining the time course of the directional effect. As with the response profile shift, the directional effect was marginally stronger at later time points. The average fractional gain for the 105 cells studied with the **4 ring/5 bar** test was 0.58 at 1 sec after target onset and 0.60 at 2 and 3 sec. This nondeclining temporal profile makes a stimulus-related effect seem unlikely.

The effect of the stimulus bar flashes, which might tend to draw involuntary attention, was addressed by examining the directional effect in the 39 cells studied with the **sustained bar** test (see above). Like the response profile shift, the directional effect was stronger in the absence of flashing probe stimuli. The average fractional gain was 0.78 in the 0–500 msec time window and 0.76 in the 500–950 msec time window. Effects were significant for 34 of 39 cells (87%) and 32 of 39 cells (82%) in the two time windows, respectively.

Another issue is the alignment of the stimuli with respect to the CRF. If the target rings were not equidistant from the true CRF center, response differences might be attributable to changes in the distance of the attentional focus. The accuracy of stimulus placement in the **4 ring/5 bar** test was assessed by finding the
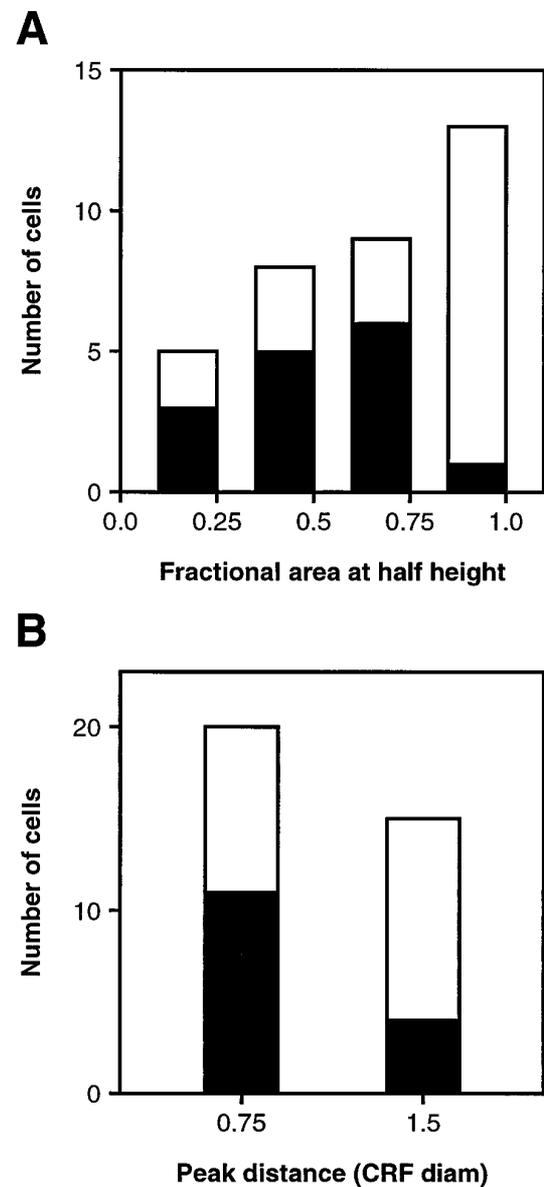

Figure 10. Directional asymmetry in the **12 ring** test (population analysis). *A,* Fractional area at half-height. The breadth of tuning for attention position was assessed by calculating the fraction of the 12 tested attention positions for which responses to the stimulus in the CRF equaled or exceeded half of the maximum response. Cells showing significant variation in response with target position according to a randomization ANOVA are shown in *black. B,* Peak distance. The spatial relationship between the peak of the attentional profile and the CRF was characterized by the distance between the target ring associated with the largest bar stimulus response and the CRF center. The **12 ring** procedure provided a test of four directions at two distances (0.75 and 1.5 RF diameters). Cells showing significant variation are shown in *black.*

spatial centers of mass for the four response profiles and averaging these as a measure of the true response center of the cell (along the axis orthogonal to bar orientation). The average absolute deviation of this value from the plotted CRF center across 105 cells was fairly small: −0.06 CRF diameters. The correlation between deviation along this axis and the fractional gain in response strength between the two target positions along the same axis was low ($r = 0.06$ signed and 0.11 unsigned) and not signifi-
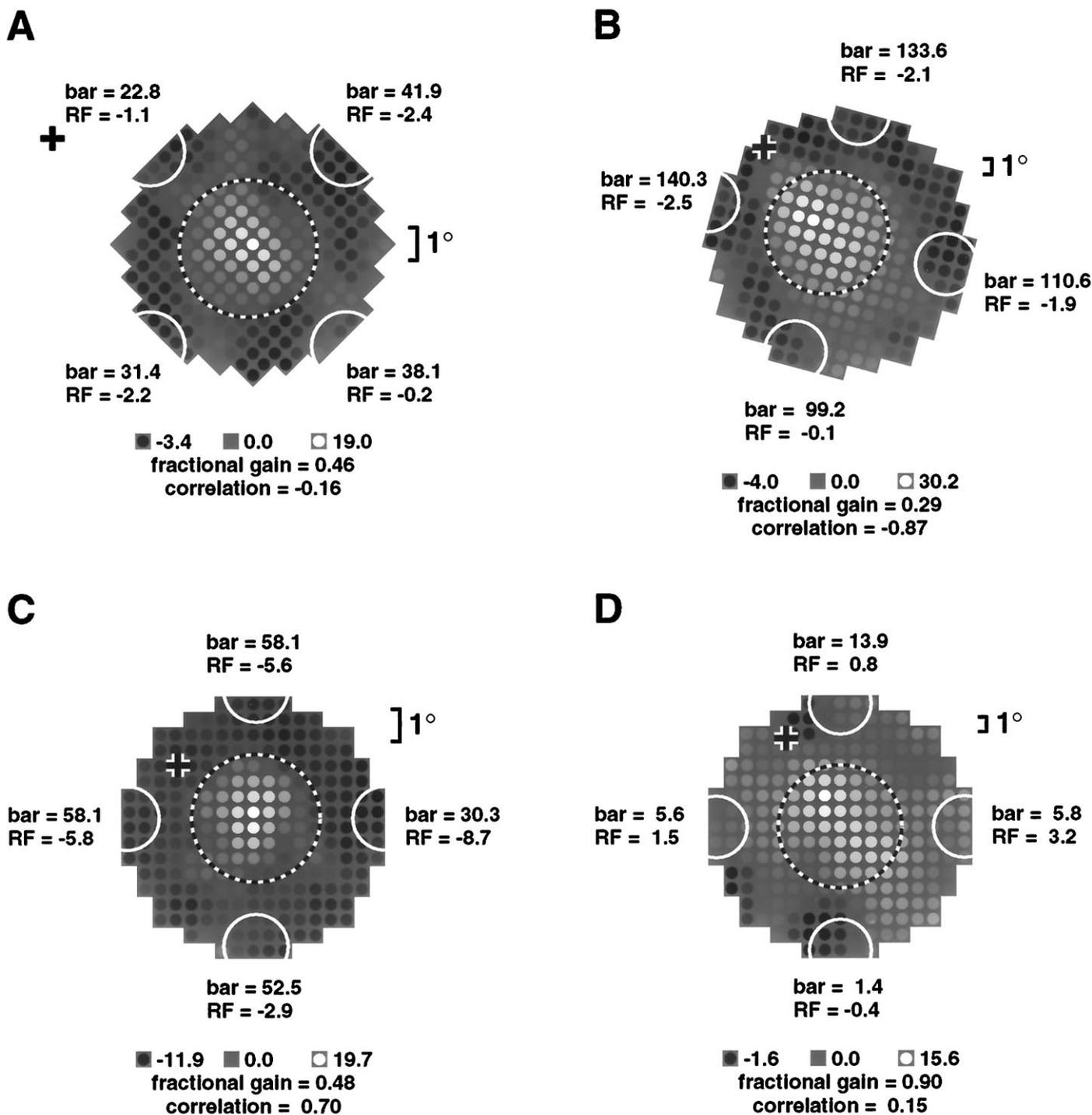
Figure 11. Comparison between receptive field structure and directional effect. In each example (*A–D*), varying gray levels are used to plot the receptive field profile for the cell. The gray levels corresponding to the minimum value (*left*), 0 (*center*), and maximum value (*right*) in spikes/sec are shown *below*. Receptive field profiles were based on an automated plotting routine that involved presentation of a small bar stimulus at each of the mapped locations (see Materials and Methods). Response rates were smoothed with a two-dimensional spatial Gaussian of SD 0.125 CRF diameters. The position of the fovea is indicated by the *cross*, and the spatial scale is given by the 1° *bar* at the *right*. The CRF, as defined for the purpose of the **4 ring/5 bar** test, is indicated by the *dashed circle*. The target rings in the **4 ring/5 bar** test are shown as *white circles*. Near each target ring are two values given in spikes/sec; "bar" indicates the summed bar response (from the **4 ring/5 bar** test) associated with attention to that ring, and "RF" indicates the average receptive field response in the region covered by the target ring and extending one-eighth of the CRF diameter beyond its boundary. The correlation between the two sets of four values is given *below*, along with the fractional gain effect for the cell.

cant ($p > 0.10$ in both cases). Thus, there is no indication that the directional asymmetry effect was based on stimulus misalignment.

The influence of attention on background firing rates also

requires consideration, because changes in background rates could produce apparent changes in bar responses. To control for this possibility, analyses were performed both with and without

background subtraction. In the **4 ring/5 bar** data set, background subtraction produced an average fractional gain of 0.55, with 89 of 105 cells (85%) showing significant effects (see above). Without background subtraction, the average fractional gain was 0.40, with 73 of 105 cells (70%) showing significant effects. This lower average value can be expected, because adding in background responses would increase absolute response rates and thus decrease proportional differences. ANOVA showed that background rate differences were significant for 23 of 105 cells (22%). For these cells, the average difference in background rate between the strongest and weakest target ring positions was 9.7 spikes/sec. Thus, background variation was sometimes significant, although it could not account for the main part of the directional attention effects.

A final issue concerning the directional effect is that of task difficulty. It has been reported previously that task difficulty can strongly influence neural responses (Spitzer et al., 1988). The possibility exists that differential difficulty in attending to the various target rings, caused for example by differences in eccentricity, might produce changes in response strength. However, error rates in this task were very low (1.5% on average, excluding fixation errors). The average correlation between error rates and response strength associated with the four target positions was 0.02. Thus, differences in task difficulty cannot explain the directional asymmetry in attention effects observed here.

DISCUSSION

The experiments presented here address a relatively unexplored issue in the neurophysiology of attention: the spatial interaction between visual stimuli and the attentional focus. The results indicate a complex modulation of responses to visual stimuli in the space immediately surrounding the attended target. First, there is a response gradient surrounding the target stimulus, such that nearby stimuli evoke stronger responses than more distant stimuli. Second, there is a differential modulation of individual cells depending on the direction of attention: some cells are more activated when attention is to the right of the stimulus, others when attention is to the left, etc. Potential mechanisms underlying these effects are discussed in the first section below. The theoretical implications of these results for spatial attention and neural coding of stimulus position are then considered.

Underlying mechanisms

Both of the effects described here involve a modulation of response strength that varies with the spatial position of the driving stimulus and the behavioral target. Both effects could be produced, therefore, by spatially varying inputs that modify the ascending sensory signals. The directional effect is analogous to the gradual modulation of parietal visual responses with changes in gaze angle (Andersen et al., 1985) and could likewise be explained by a modulatory gain field, in this case dependent on the position of attention rather than the eye. The shift effect is consistent with the notion of a spatial "spotlight" of relative firing rate enhancement imposed on retinotopically mapped visual areas (cf. Crick, 1984; Koch and Ullman, 1985; Tsotsos et al., 1995).

Alternatively, the shift effect may reflect an actual translation of the receptive field, by means of a dynamic remapping of inputs from lower levels (Olshausen et al., 1993). This shifter circuit model predicts that progressive shifts in response profiles at successive stages in the visual hierarchy lead to large-scale translations at higher stages, thereby preserving spatial relationships within the attentional focus. Thus, determining whether larger

shifts occur in inferotemporal cortex will constitute an important test of this model.

Another relevant proposal is the idea that receptive fields in V4 and other areas shrink around the behavioral target object (Moran and Desimone, 1985). The results reported here are inconsistent with wholesale receptive field shrinkage, because the response profiles generally matched the plotted CRF in size (see Fig. 4). Nonetheless, the shift effect is compatible with the spatial biasing of response strength reported by Moran and Desimone, because a shifting spatial profile of enhancement centered on the target object could account for their results. Moran and Desimone also reported differential effects of directing attention within versus outside the CRF. Our experiments do not speak directly to this point, because in all of our conditions attention was directed outside the CRF. However, our findings clearly show that changing the position of attention outside the CRF affects responses to stimuli within the CRF.

The possibility that the directional effect is based on asymmetries in the receptive field surround was considered in Results. Our analyses, based on automated receptive field plots using single bar stimuli, provide no support for this notion. It remains possible that a relationship would be revealed by experiments testing explicitly for suppressive sensory interactions between stimuli simultaneously presented inside and outside the CRF. However, the existence of such a relationship would not alter the attentional nature of the directional effect reported here, which is attributable to changes in behavioral state, not stimulus changes (see Results). It would mean merely that attention acts by determining which surround stimulus influences responses [cf. Desimone and Duncan (1995), arguing that attention is primarily a mechanism for biasing competition between multiple stimuli]. Moreover, the functional significance of the effect would remain the same: in complex visual situations with multiple stimuli in the surround, V4 responses would depend on the positional relationship between stimulus and attention.

Spatial attention

Psychological studies show that visual attention has a strong spatial aspect: perceptual processing tends to be enhanced within a spatial region centered on the attentional cue or target (Eriksen and Eriksen, 1974; Posner et al., 1980; Hoffman and Nelson, 1981). Some results suggest that the window is variable in size (Eriksen and Yeh, 1985) and falls off gradually with distance (Downing and Pinker, 1985). The results presented here provide the first direct neurophysiological correlate for a spatial window of attention. Responses to stimuli in close proximity to the attentional target were enhanced relative to responses to more distant stimuli (Figs. 2–4). Effects were larger at positions closer to the attentional focus (positions *1* and *5* vs positions *2* and *4* in Fig. 4), consistent with psychological demonstrations of a spatial gradient of attentional enhancement. The gradient of bar response modulation extended at least 0.5 CRF diameters beyond the boundary of the target object (to bar positions *2* and *4*). This was the effective limit of the tested region, because more distant bar positions were equidistant or closer to the comparison target position (see Fig. 1*A*). The corresponding absolute distance varied from 1.1° to 5.0°, with an average of 3.1° (depending on CRF diameter). This spatial range is consistent with psychological measurements showing a gradient that extends at least 5° to 6° from the target (Downing and Pinker, 1985).

Space is not the only dimension for allocating attention. A major conceptual alternative is that the visual world is preatten-

tively segmented into objects and that attention is subsequently allocated to one of these objects, rather than to a region of space (Duncan, 1984). Psychological evidence supports the existence of both spatial and object-based attention under different circumstances, and the results presented here do nothing to exclude the possibility of object-based mechanisms. Attention can also be allocated along other feature dimensions, such as color, and Motter (1994a,b) has provided striking evidence for nonspatial modulation of neural responses based on attention to color.

Position coding

We propose that the directional asymmetry in spatial attention effects described above reflects a mechanism for encoding the positions of visual features relative to the attentional focus. Feature position is initially represented in retinotopic coordinates, but an explicit representation of local relationships ("above," "left of," etc.) would be more efficient for many purposes (Kosslyn et al., 1992). In particular, a local representation of this sort could achieve invariance to translation and scaling, a key requirement for pattern recognition systems. Attention would presumably be required to select a particular region out of all of the potential local regions in a natural scene. Indeed, several results indicate that perception of local position relationships requires attention (Logan, 1994; Wolfe and Bennett, 1997). The V4 neurons described here would provide an attention-based, explicit representation of local position relationships. Some cells are tuned for stimuli above the currently attended object, some are tuned for stimuli to the left, etc. (Figs. 5, 6, 9). The breadth of tuning seen for individual cells suggests that these positional relationships are encoded by the combined activity of cell populations.

The local position information provided by these cells would be useful for parsing the spatial structure of multi-part objects, such as faces. A face is distinguishable not only by the presence of certain features but also by their relative positions: the mouth near the bottom, the nose in the middle, the eyes near the top, etc. Figure 12 illustrates how cells with the directional tuning properties described here could represent feature positions. With attention centered on the face, the mouth would drive cells selective for horizontal stimuli that lie below the center of attention (i.e., cells that respond best when attention is above the stimulus). Cells selective for horizontal stimuli that lie above (or to the left or right of) the center of attention would not be driven. Activation of "horizontal/below" cells would provide part of the combined feature/position information necessary for perceiving the face. This information would be available across a range of fixation locations, as long as attention remained roughly centered on the face. Large changes in fixation would bring other cells into play as the retinotopic location of the mouth changed, but these cells too would be tuned for position relative to attention. At higher levels, such as the more anterior parts of IT, classical receptive fields are larger and a single group of cells would suffice. With such large receptive fields, there would be a loss of retinotopic position coding, and a mechanism like this would be useful for preserving information about positional relationships (cf. Olshausen et al., 1993).

The ultimate goal of the mechanism described above is a transformation from retinotopic coordinates into a reference frame centered on the currently attended object (cf. Hinton and Parsons, 1981). Such a transformation is implicit in the responses seen here, and computational modeling shows that an invariant representation of spatial relationships can be extracted from these responses (Salinas and Abbott, 1997). This type of implicit representation would be analogous to the implicit body-centered
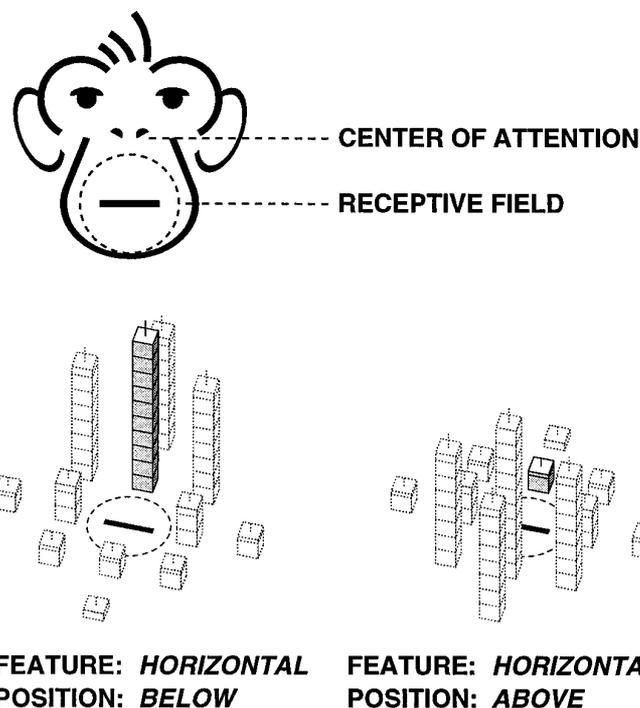

FEATURE: *HORIZONTAL* POSITION: *BELOW*   **FEATURE: *HORIZONTAL* POSITION: *ABOVE***

Figure 12. Proposed strategy for local position coding. Attention is centered on the face (for example, around the nose or eyes), although the window of attention would encompass the entire head. A V4 receptive field (*dashed outline*) is shown at the position of the mouth. A cell with this receptive field location that was sensitive to horizontal stimuli and had the attention tuning profile on the left would respond well under these circumstances, because the position of attention (*shaded block*) is at the peak of the tuning profile. A cell with the tuning profile on the right would respond poorly. The selective activation of the cell on the left would represent the presence of a horizontal stimulus below the center of attention and thus contribute to perception of the face. The tuning profiles for the two cells are idealized versions in the format of Figure 9. Block height represents average response to horizontal stimuli in the CRF (*dashed circle*) when attention is centered on the corresponding location.

representation in parietal cortex (Andersen et al., 1985; Zipser and Andersen, 1988). Alternatively, the responses observed in V4 may represent the intermediate stages of a transformation that would be realized explicitly in IT by cells with dynamic receptive fields defined in attention-centered coordinates (Olshausen et al., 1993). In either case, the effects shown here would represent a neural analog to object-centered psychological phenomena (Tipper et al., 1991; Driver et al., 1992; Baylis and Driver, 1993; Halligan and Marshall, 1993; Gibson and Egeth, 1994) (cf. Olson and Gettner, 1995).

REFERENCES

Allman J, Miezin F, McGuinness E (1985) Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. Annu Rev Neurosci 8:407–430.

Andersen RA, Essick GK, Siegel RM (1985) Encoding of spatial location by posterior parietal neurons. Science 230:456–458.

Baylis GC, Driver J (1993) Visual attention and objects: evidence for hierarchical coding of location. J Exp Psychol 19:451–470.

Connor CE, Gallant JL, Preddie DC, Van Essen DC (1996) Responses in area V4 depend on the spatial relationship between stimulus and attention. J Neurophysiol 75:1306–1308.

Crick F (1984) Function of the thalamic reticular complex: the searchlight hypothesis. Proc Natl Acad Sci USA 81:4586–4590.

Crick F, Koch C (1990) Towards a neurobiological theory of consciousness. Semin Neurosci 2:263–275.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

Desimone R, Schein SJ, Moran J, Ungerleider LG (1985) Contour, color and shape analysis beyond the striate cortex. Vision Res 3:441–452.

Downing CJ, Pinker S (1985) The spatial structure of visual attention. In: Mechanisms of attention: attention and performance XI (Posner MI, Marin OSM, eds), pp 171–187. Hillsdale, NJ: LEA.

Driver J, Baylis GC, Rafal RD (1992) Preserved figure-ground segregation and symmetry perception in visual neglect. Nature 360:73–75.

Duncan J (1984) Selective attention and the organization of visual information. J Exp Psychol 113:501–517.

Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept Psychophysiol 16:143–149.

Eriksen CW, Yeh Y (1985) Allocation of attention in the visual field. J Exp Psychol 11:583–597.

Gibson BS, Egeth H (1994) Inhibition of return to object-based and environment-based locations. Percept Psychophysiol 55:323–339.

Halligan PW, Marshall JC (1993) When two is one: a case study of spatial parsing in visual neglect. Perception 22:309–312.

Hinton GE, Parsons LM (1981) Frames of reference and mental imagery. In: Attention and performance IX (Long J, Baddely A, eds), pp 261–277. Hillsdale, NJ: LEA.

Hoffman JE, Nelson B (1981) Spatial selectivity in visual search. Percept Psychophysiol 30:283–290.

Knierim JJ, Van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol 67:961–980.

Koch C, Ullman S (1985) Shifts in selective visual attention: towards the underlying neural circuitry. Hum Neurobiol 4:219–227.

Kosslyn SM, Chabris CF, Marsolek CJ, Koenig O (1992) Categorical versus coordinate spatial relations: computational analyses and computer simulations. J Exp Psychol 18:562–577.

Logan GD (1994) Spatial attention and the apprehension of spatial relations. J Exp Psychol 20:1015–1036.

Manly BFJ (1991) Randomization and Monte Carlo methods in biology. New York: Chapman and Hall.

Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229:782–784.

Motter BC (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol 70:1–11.

Motter BC (1994a) Neural correlates of attentive selection for color or luminance in extrastriate area V4. J Neurosci 14:2178–2189.

Motter BC (1994b) Neural correlates of feature selective memory and pop-out in extrastriate area V4. J Neurosci 14:2190–2199.

Niebur E, Koch C (1994) A model for the neuronal implementation of selective visual attention based on temporal correlation among neurons. J Comput Neurosci 1:141–158.

Olshausen BA, Anderson CH, Van Essen DC (1993) A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. J Neurosci 13:4700–4719.

Olson CR, Gettner SN (1995) Object-centered direction selectivity in the macaque supplementary eye field. Science 269:985–988.

Posner MI, Snyder CRR, Davidson BJ (1980) Attention and the detection of signals. J Exp Psychol 109:160–174.

Robinson DA (1963) A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE Trans Biomed Eng 10:137–145.

Salinas LE, Abbott LF (1997) Invariant visual perception from attentional gain fields. J Neurophysiol, in press.

Spitzer H, Desimone R, Moran J (1988) Increased attention enhances both behavioral and neuronal performance. Science 240:338–340.

Tipper SP, Driver J, Weaver B (1991) Object-centred inhibition of return of visual attention. Q J Exp Psychol 43A:289–298.

Tsotsos JK, Culhane SM, Kei Wai WY, Lai Y, Davis N, Nuflo F (1995) Modeling visual attention via selective tuning. Artificial Intelligence 78:507–547.

von der Malsburg C, Schneider W (1986) A neural cocktail party processor. Biol Cybern 54:29–40.

Wolfe JM, Bennett SC (1997) Preattentive object files: shapeless bundles of basic features. Vision Res, in press.

Yantis S (1993) Stimulus-driven attentional capture. Curr Dir Psychol Sci 2:156–161.

Zipser D, Andersen RA (1988) A back-propagation programmed network that simulated response properties of a subset of posterior parietal neurons. Nature 331:697–684.