

# Effects of Attention on MT and MST Neuronal Activity During Pursuit Initiation

GREGG H. RECANZONE<sup>1</sup> AND ROBERT H. WURTZ<sup>2</sup>

<sup>1</sup>*Center for Neuroscience and Section of Neurobiology, Physiology and Behavior, University of California at Davis, Davis, California 95616; and* <sup>2</sup>*Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892*

**Recanzone, Gregg H. and Robert H. Wurtz.** Effects of attention on MT and MST neuronal activity during pursuit initiation. *J. Neurophysiol.* 83: 777–790, 2000. The responses of neurons in monkey extrastriate areas MT (middle temporal) and MST (medial superior temporal), and the initial metrics of saccadic and pursuit eye movements, have previously been shown to be better predicted by vector averaging or winner-take-all models depending on the stimulus conditions. To investigate the potential influences of attention on the neuronal activity, we measured the responses of single MT and MST neurons under identical stimulus conditions when one of two moving stimuli was the target for a pursuit eye movement. We found the greatest attentional modulation across neurons when two stimuli moved through the receptive field (RF) of the neuron and the stimulus motion was initiated at least 450 ms before reaching the center of the RF. These conditions were the same as those in which a winner-take-all model better predicted both the eye movements and the underlying neuronal activity. The modulation was almost always an increase of activity, and it was about equally frequent in MT and MST. A modulation of >50% was observed in ~41% of MT neurons and 27% of MST neurons. Responses to all directions of motion were modulated so that the direction tuning curves in the attended and unattended conditions were similar. Changes in the background activity with target selection were small and unlikely to account for the observed attentional modulation. In contrast, there was little change in the neuronal response with attention when the stimulus reached the RF center 150 ms after motion onset, which was also the condition in which the vector average model better predicted the initial eye movements and the activity of the neurons. These results are consistent with a competition model of attention in which top-down attention acts on the activity of one of two competing populations of neurons activated by the bottom-up input from peripheral stimuli. They suggest that there is a minimal separation of the populations necessary before attention can act on one population, similar to that required to produce a winner-take-all mode of behavior in pursuit initiation. The present experiments also suggest that it takes several hundred milliseconds to develop this top-down attention effect.

## INTRODUCTION

The direction and speed of moving visual stimuli are encoded in extrastriate cortical areas MT (middle temporal area) and MST (medial superior temporal area) (Allman et al. 1985; Desimone and Ungerleider 1986; Maunsell and Van Essen 1983; Tanaka et al. 1989; Van Essen et al. 1981). These cortical areas are also instrumental in the generation of smooth pursuit eye movements, based on both the activity of the

neurons during this behavior and deficits following ablation (Dursteler et al. 1987; Groh et al. 1997; Komatsu and Wurtz 1988, 1989; Newsome et al. 1985; Schiller and Lee 1994; Thier and Erickson 1992; Yamasaki and Wurtz 1990). The results of our previous experiments indicated that during a task in which the monkey was required to pursue one of two stimuli, the responses of MT and MST neurons could be better represented by one of two models depending on the stimulus conditions (Recanzone et al. 1997; Recanzone and Wurtz 1999). The model that better described the metrics of the initial eye movements and the activity of the MT and MST neurons was a vector average when both the time between the onset of target motion and pursuit initiation (150 ms) and the distance over which the target moved was short. A winner-take-all model was the better description when the time (450 ms) and distance was longer (Recanzone and Wurtz 1999). This difference was interpreted as a reflection of the interactions of subpopulations of neurons within MT and MST that have receptive fields and best directions of motion aligned along the stimulus trajectory.

In these previous experiments, we used the motion of single stimuli in the receptive field (RF) to predict the response when two stimuli moved through the RF simultaneously. On all trials two different stimuli moved through the visual field, and the monkey presumably directed its attention to the target stimulus but not to the other stimulus. In the present experiments we have specifically investigated the contribution of attention to the responses of MT and MST neurons by comparing their responses on trials when we required the monkey to attend to one stimulus with the responses on trials when we required the monkey to attend to the other stimulus. We did this attention shift by requiring the monkey to use one stimulus or the other as the target for a pursuit eye movement. We compared the neuronal responses on long and short-duration trials that led to winner-take-all and vector average interpretations of the behavior and the underlying neuronal activity, and also when two stimuli moved through the RF of the neuron and when the two stimuli moved through different visual hemifields. In an additional set of experiments, we changed the cue for attention from the usual shape of the fixation stimulus to one of the locations in the visual field. We found only limited attentional modulation under conditions that were better predicted by the vector average model, and the largest attentional modulation under conditions that were better predicted by the winner-take-all model, even for the same neurons tested on interleaved trials. These attentional effects were intermediate in magnitude compared with two recent reports of attention effects in these

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.

visual motion processing areas (Seidemann and Newsome 1999; Treue and Maunsell 1999). In the present experiments we were able to show attentional effects that were restricted to certain conditions on randomly interleaved trials, and we believe the results offer a few clues to the conditions under which attention effects are most prominent.

A brief report of this study has appeared previously (Recanzone et al. 1993).

## METHODS

The general procedures in these experiments were identical to those described in detail in Recanzone and Wurtz (1999), and here we will provide those methods that are critical for understanding the current attention experiments. All procedures were approved by the Institute Animal Care and Use Committee and complied with Public Health Service policy on the humane care and use of laboratory animals.

### Target selection task

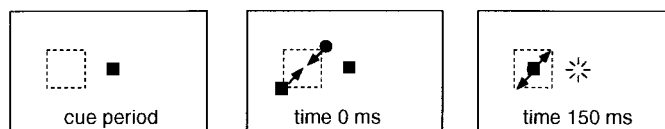
Two adult male rhesus monkeys (*Macaca mulatta*) were used in this study (*monkeys N and P*). Each monkey was trained to perform several versions of two pursuit eye movement tasks, and these tasks are illustrated in Fig. 1. In all versions of the task, there were three steps that are shown in the three columns in Fig. 1. First, the monkey was required to fixate a central stimulus (middle square in *left panels* of Fig. 1) to within  $\pm 2^\circ$  for a random period of 100–500 ms. Second, two stimuli were presented in motion (Fig. 1, *middle panels*), but the monkeys were required to maintain fixation for 150 or 450 ms. Third, the fixation stimulus was extinguished (Fig. 1, *right panels*), and the monkey had to make a saccade to the location of the target and then match the eye movement to the target motion with smooth pursuit to receive a liquid reward. The stimulus conditions were set so that at least one of these two stimuli always moved through the RF of the neuron under study (dashed box in Fig. 1) and so that this stimulus was at the geometric center of the RF when the monkey was released from fixation.

The cue used in the first set of experiments was the shape of the fixation stimulus. In this shape-cue task, the monkey was instructed which stimulus was to be the target on that trial by the shape of the fixation stimulus that matched the shape of one of the forthcoming targets. For example, in Fig. 1A the square fixation stimulus matched the upward moving target so that correct pursuit was to that target. On other trials the fixation stimulus was a circle, and the monkey was thereby instructed to pursue the downward moving target.

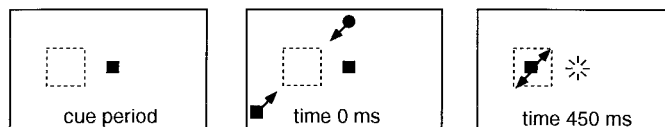
The timing and location of the onset of the stimulus motion was varied. On the short-duration trials shown in Fig. 1A, the stimulus crossed the RF in 150 ms after the start of motion. On the long-duration trials the moving stimuli were displaced in space at the onset of the trial and therefore had a longer trajectory to the center of the RF, but moved at the same velocity. These stimuli reached the center of the RF 450 ms after moving stimulus onset. We term these trial types as “long duration” and “short duration,” although it must be emphasized that both the time of motion and the space over which the motion occurred varied between the two stimulus conditions.

The location of the onset of the stimulus motion was also varied, although one always moved through the RF of the neuron under study. On half the trials the other stimulus moved in the same visual hemifield as the RF stimulus and the two stimuli intersected at the RF center. We refer to these as “crossed” trials, and they are illustrated in Fig. 1, A and B (the same convention as used in the previous paper, Recanzone and Wurtz 1999). On the other half of the trials, the second stimulus moved in the opposite visual hemifield and did not cross the path of the RF stimulus, and we refer to trials with this stimulus configuration as “uncrossed” trials, as shown in Fig. 1C. Both short-duration and

### A Shape - Cue Task, Crossed Stimuli, Short-Duration



### B Shape - Cue Task, Crossed Stimuli, Long-Duration



### C Shape - Cue Task, Uncrossed Stimuli, Long-Duration



### D Location - Cue Task, Long-Duration

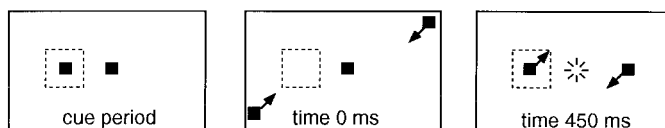


FIG. 1. Schematic diagram of the behavioral paradigms. Representative stimulus configurations at 3 different time points are illustrated for most variations of the behavioral tasks used. *Left panels*: visual display prior to the monkey fixating the central stimulus (■). The large rectangle represents the tangent screen in front of the monkey, and the dashed box represents the location of the receptive field of the neuron under study. *Middle panels*: visual display on the frame where the 1st moving stimuli were presented. *Right panels*: visual display either 150 ms (A) or 450 ms (B–D) after moving stimulus onset, when the monkey was released from fixation and required to make a saccadic and smooth pursuit eye movement toward the target. A: shape cue task crossed trial with short-duration where the target (■) moved in the best direction through the receptive field of the neuron at the same time that the other stimulus (●) also moved through the receptive field. B: long-duration version of the same task. In this case, the stimuli started their motion farther from the receptive field and did not intersect at the center until 450 ms had elapsed from motion onset. C: uncrossed trials with long duration for the same task. In this case, each stimulus was presented in a different hemifield. For A–C the target was the square moving up and to the right on each trial (“attended” trials), and the same moving stimulus configuration was also presented when the fixation stimulus was a circle, indicating that the circle was the target stimulus on those trials (“unattended” trials). D: location cue long duration trial. In this case, a cue was presented in the visual hemifield where the target would be located prior to the monkey fixating the fixation stimulus. All visual stimuli are the same shape, and only the preceding cue indicated which was the target. RF, receptive field of a hypothetical neuron. Arrows indicate the direction of motion of each of the 2 shapes.

long-duration trial types were included in both the crossed and uncrossed trials.

Finally, we used trials with a different cue condition, a location cue. On these trials the monkey was cued as to which visual hemifield the target stimulus would be presented in (Fig. 1D). On this task, all visual stimuli including the fixation stimulus were the same shape, and the moving stimuli were always in different visual hemifields. Short-duration and long-duration trial types were included in the location cue task as well, and of course all of these trials were of the uncrossed type.

We will operationally define the “attended” trials as those in which the monkey selected and made an eye movement to a stimulus passing through the RF and the “unattended” trials as those in which the eye movement was made to the other stimulus. On all trials there were two

moving stimuli, and the stimuli in the RF were physically identical between attended and unattended trials. For the uncrossed trials, those in which the stimuli moving through the RF were the targets for the eye movements are termed the attended trials, whereas those in which the identical stimuli passed through the RF but the targets were the stimuli in the opposite visual field are termed the unattended trials. For the crossed trials, the two stimuli always moved through the RF, one in the best direction and one in a nonbest direction. Trials in which the targets moved in the best direction are termed the attended trials, whereas trials in which the target stimuli moved in a nonbest direction while the other stimulus moved in the best direction are termed the unattended trials.

Thus the present analysis compared the neuronal responses to the same stimulus on the attended and unattended trials. In contrast, the previous experiments (Recanzone and Wurtz 1999) compared the neuronal response with one stimulus in the RF on uncrossed trials to the case with two stimuli in the RF on crossed trials to determine whether the integration was better described by an average or a winner-take-all model. In that previous analysis, the attention to the stimuli was the same on all the analyzed trials.

Stimuli moved in eight different directions (0, 45, 90, 135, 180, 225, 270, and 315°). For shape cue and location cue uncrossed trials, the stimulus that was the target moved in any of the eight directions on any trial, as could the stimulus in the other visual hemifield. For the shape-cue crossed trials, one stimulus always moved in the best direction, and the other stimulus moved in one of the seven other nonbest directions. On these trials, either stimulus could be the target for the eye movement. In a given session, only location cue trials or shape cue trials were presented, but within these sessions all of the stimulus conditions (crossed trials and uncrossed trials, different target directions) were presented on randomly interleaved trials, and when long-duration trials were used, they also were introduced with the other trial types in randomly interleaved order. A complete data set consisted of at least eight correct trials for each stimulus type, although most of the data of this report are from 10–15 correct trials for each stimulus type. The majority of neurons were recorded in both the shape-cue and location-cue types of sessions, although these could be initiated up to 90 min apart in time. In some instances, the neuron was lost before a complete data set could be recorded under both experimental conditions. The results of the behavioral measures (saccadic and pursuit eye movements) and the behavioral performance have been reported previously (Recanzone and Wurtz 1999).

For the shape-cue task, *monkey N* completed 80 sessions and *monkey P* completed 65 sessions with short-duration stimuli in which both crossed and uncrossed trials were presented (Fig. 1, A–C). For the location-cue task, *monkey N* completed 92 sessions and *monkey P* completed 55 sessions. For the long-duration stimuli, *monkey N* completed 60 sessions for the shape-cue task and 72 sessions for the location-cue task, but *monkey P* frequently was unable to reliably suppress the initial saccade for a sufficient period to complete at least 8 trials under all conditions for the long-duration trials (only 13 and 7 sessions with only 3–7 successful trials per stimulus per session for the location-cue and shape-cue task, respectively). Although we found that these limited data were consistent with those obtained in *monkey N*, this data set is not sufficiently extensive to report in detail in results.

### Experimental procedures

Visual stimuli were back-projected onto a tangent screen placed 57.4 cm from the monkey using a video projector (Electrochrome, SVGA, 1,024 × 768 pixel resolution). Each pixel subtended a visual angle of 0.13° horizontally and 0.12° vertically. Images were created by a PC and were presented at a frame rate of 72 Hz. Stimulus objects were brighter (1.8 cd/m<sup>2</sup>) than the background (0.2 cd/m<sup>2</sup>). Two of five different objects (circle, square, diamond, plus sign, and triangle) were used in each behavioral session, with each object of equal

luminance and size (same number of pixels/object) subtending a maximum visual angle of 1.8°. Objects were moved across the screen by displacing each illuminated pixel by 1 or 2 pixels each frame in either the horizontal, vertical, or both directions. Three stimulus speeds were produced by pixel shifts of 1, 1.5, or 2 pixels/frame corresponding to ~9, 13.5, or 18°/s along the horizontal and vertical directions, and 12, 18.5, or 25°/s along the obliques. For speeds using 1.5 pixels/frame, all stimulus pixels were displaced by 1 and 2 pixels on alternate frames. These stimulus motions were perceived by human observers as continuous, smooth motion.

Monkeys underwent magnetic resonance imaging (MRI) scans in the frontal and sagittal planes (1.5 T magnet, 3 mm slices) before implantation of a restraining head post and scleral search coils in each eye. The MRI images allowed us to direct electrode penetrations into the region of the superior temporal sulcus just medial to the opening of the sulcus into a floor. Following initial training, a recording cylinder was implanted directly over areas MT and MST in the stereotaxic vertical plane (see Recanzone and Wurtz 1999). Single neuron activity was recorded in cortical areas MT and MST, whereas horizontal and vertical eye position was recorded using the magnetic search coil technique (Fuchs and Robinson 1966; Judge et al. 1980). Both eye position and unit activity were monitored on-line as well as stored for off-line analysis. Neurons recorded in this study met the same three criteria described previously (Recanzone et al. 1997; Recanzone and Wurtz 1999): 1) the activity of the neuron was altered by the presence of visual stimuli (but not necessarily moving stimuli), 2) isolation was sufficient to be confident that only a single neuron was being recorded, and 3) the center of the RF was between 5 and 25° of eccentricity. This last requirement was necessary because the monkeys had difficulty in maintaining fixation if a stimulus moved through the fixation stimulus before the fixation stimulus offset and reliably making the visual shape discrimination at eccentricities >25°.

RFs were defined using hand-held bars, spots, and patterns of light. RF edges were defined as locations where the neuron no longer responded to moving, flashed, or stationary visual stimuli. We categorized each neuron as being located within cortical area MT or MST based on the location and depth of the electrode within the recording cylinder relative to the MRI images, and the characteristics of the visual stimuli required to maximally activate the neuron (Desimone and Ungerleider 1986; Maunsell and Van Essen 1983; Van Essen et al. 1981). We only rarely encountered neurons with very large RFs that responded best to large patterned stimulation similar to those described in the dorsal region of MST (MSTd) (Komatsu and Wurtz 1988; Tanaka and Saito 1989; Tanaka et al. 1986, 1989, 1993), and the vast majority of our sample of MST neurons were characteristic of those located in the lateral region (MSTl). However, it is possible that some of the neurons in this study were located within MSTd. Histological sections through the superior temporal sulcus obtained at the end of the experiments showed that the electrode tracks were consistent with our categorization of neurons into MT or MST, as described previously (Recanzone et al. 1997).

The stimulus speed and shape used in a particular behavioral session was based on the best response of the recorded neuron during preliminary characterization of the cell's tuning and RF properties and was adjusted so that the stimulus would traverse as much of the RF as possible before reaching the center on the short-duration trials. RF sizes ranged from diameters of 5–24° for neurons recorded in MT, and 5–35° for neurons recorded in MST, and increased in diameter with increasing eccentricity, as described previously (e.g., Desimone and Ungerleider 1986; Komatsu and Wurtz 1988). Stimulus parameters used were tailored to each neuron under study such that the onset of the motion stimulus occurred near the edge of the RF. We noted no differences in the response properties of these neurons under the different stimulus conditions when the stimulus onset was slightly within, at, or slightly outside the edge of the RF, and thus all data were pooled.

### Data analysis

For the responses of individual MT and MST neurons reported here, the firing rate was measured during the period from the onset of the moving visual stimuli to the onset of the initial saccade for all correct trials. The firing rate before the onset of the moving stimulus was also calculated (spontaneous activity) and subtracted from the driven activity. For each neuron, the best direction was defined as the stimulus direction that gave the largest response during the uncrossed trials. The null direction was defined as the direction 180° from the best direction. Given the low occurrence of incorrect trials, we were not able to directly compare the responses of correct and incorrect trials.

We used an attention index that was a contrast ratio calculated following the procedures of Treue and Maunsell (1996). The attention index was defined as the difference in the average response between attended and unattended trials divided by the sum of the response. This value varies from  $-1.0$  (no attended response) to  $+1.0$  (no unattended response), with zero where there is no difference between the responses under the two conditions.

The effect of attention on the response to different directions of motion was defined by the vector strength. For the uncrossed trials, a vector was defined for each direction of motion with the vector direction equal to the direction of the target motion and the amplitude of the vector as the neural response (mean spikes/trial). These vectors were then summed and divided by the total number of spikes. This metric varies from zero (no directional tuning) to one (all responses for only one direction). For the crossed trial condition, because one stimulus always moved in the best direction of the neuron, only seven directions of motion could be used in calculating the vector strength. For this analysis, the attended vector strength was calculated from trials in which the target moved in the best direction, and the other stimulus moved in each of the seven nonbest directions. The unattended vector strength was calculated from trials where the nontarget stimulus moved in the best direction and the target stimulus moved in one of seven nonbest directions. To quantify differences between the vector strengths on the attended and unattended trials we used a contrast ratio with the same calculation as for the attention index described above.

## RESULTS

### Attentional modulation during the shape-cue task

We determined the effect of attending to one of two stimuli moving across the RF on a total of 107 MT neurons and 105 MST neurons in the two monkeys. Figure 2 shows the rasters and peristimulus time histograms from a typical MT neuron during trials in which the two stimuli both crossed the RF and in which the stimulus attended to was indicated by the shape of the fixation stimulus (shape-cue task with crossed stimuli; Fig. 1, *A* and *B*). In the *left column* of Fig. 2 are the attended trials in which the target for the eye movement was moving across the RF of the neuron in its best direction (leftward moving stimulus, indicated to the monkey by the target being a square like the fixation stimulus). In the *right column* are the unattended trials in which the target for the eye movement moved in the null direction of the neuron (rightward moving stimulus, the fixation stimulus was a circle on these trials). Note that on both the attended and unattended trials the stimuli crossing the RF of the neuron were identical: the only difference in the visual stimulus was the shape of the fixation point, and the only difference to the monkey was which stimulus was the target for the eye movement on a given trial. The response of this MT neuron showed very little difference between the attended and

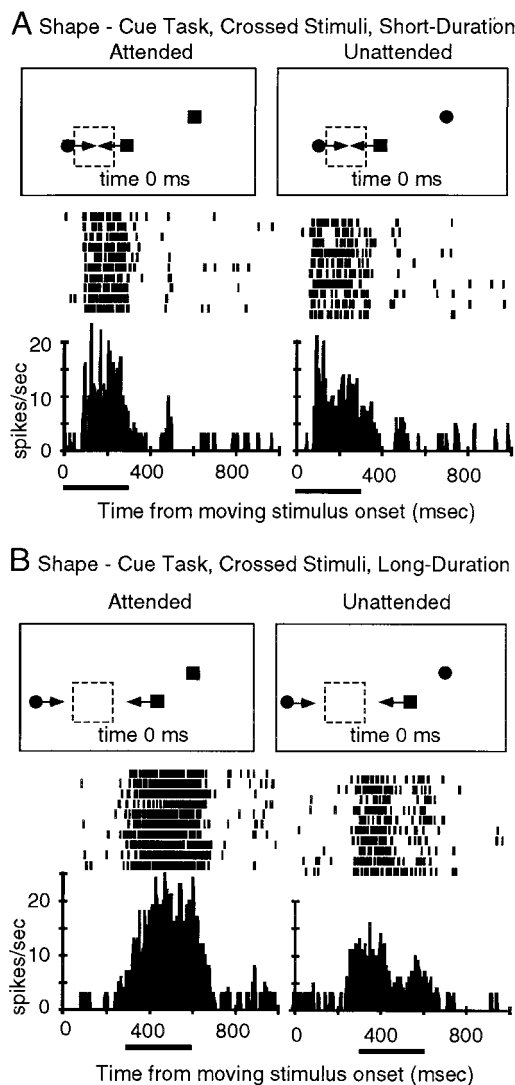


FIG. 2. Single neuron responses on crossed trials during the shape-cue task. *A*: rasters and peristimulus time histograms (PSTHs) during the short-duration trials when the target moved in the best direction of the neuron under study (*left*) and in the null direction (*right*) on randomly interleaved trials. Each tic represents a single action potential, each line a single trial. PSTH binwidth is 3 ms. Heavy horizontal bar indicates the approximate time that the stimulus was in the RF of the neuron. For these trials, the attention index ( $[(\text{best} - \text{null})/(\text{best} + \text{null})]$ ) was  $-0.01$ . *B*: rasters and PSTHs during the long duration trials. The attention index was 0.37. *Inset*: stimulus configuration on the 1st frame in which the moving stimuli were presented; stimulus conditions between the *left* and *right* panels were identical except for the shape of the fixation stimulus, which was outside of the receptive field. Dashed box: receptive field. Only the long-duration trials showed any difference in activity depending on which stimulus was the target.

unattended trials when there was only a short temporal and spatial interval between the stimulus motion onset and the time that the stimulus intersected at the RF center (150 ms; short-duration trials, Fig. 2*A*), and this difference was not statistically significant (2-tailed, unpaired *t*-test;  $P > 0.05$ ). In contrast, by separating the two stimuli in time and space between the time of stimulus motion onset and crossing the RF (450 ms; long-duration trials, Fig. 2*B*), there was clearly a stronger response when the stimulus moving in the best direction was attended (*left panel*) than when it was not (*right panel*). There was a

statistically significant difference in the response between these two trial types ( $P < 0.05$ ).

Across the population of MT neurons, 12/72 (17%) showed statistically significant differences in activity as a function of attention on the short-duration crossed trials (2-tailed unpaired  $t$ -test;  $P < 0.05$ ). These neurons were equally divided between those that showed statistically significant increases (6 neurons) and decreases (6 neurons). In MST, there were 14/73 (19%) neurons showing statistically significant modulation, with 11% showing increases and 8% showing decreases. For the long-duration trials, however, there were 14/27 neurons (52%) in MT that showed statistically significant differences, with only one neuron showing a decrease in activity as a function of attention and the remaining 13 all showing increases. Similarly, 12/33 (36%) of MST neurons had statistically significant modulation, and all neurons showed increases with attention. The median increase in the response on the long-duration trials was 11% for MT neurons and 16% for MST neurons.

To quantify the size of the effect of attention on each neuron, we used a contrast ratio calculated as the difference between the attended and unattended trials divided by their sum. Positive values of this index indicate an increased response with attention, negative values indicate a reduced response, and zero indicates no change. Figure 3 shows the distribution of attention indexes. In both MT and MST (Fig. 3A), for the short-duration trials of the shape-cue task there were approximately equal numbers of positive and negative values, and the distribution of attention indexes across the sample was not significantly different from a population with a mean of zero ( $P > 0.05$ ). Comparison between MT and MST similarly showed no statistically significant difference (2-tailed, unpaired  $t$ -test;  $P > 0.05$ ). In contrast, for the long-duration trials (Fig. 3B), in both MT and MST the sample mean was significantly greater than zero ( $P < 0.05$ ) between these trial types, with a mean of 0.17 for MT neurons and 0.15 for MST neurons. Comparison between MT and MST again showed no statistically significant difference ( $P > 0.05$ ). Although the changes on the long-duration trials were significant, they were generally small. A doubling of the response (100% increase), which was approximately the median for these two areas in a previous report (86 and 113% for MT and MST, respectively) (Treue and Maunsell 1996) was only observed in 19% (5/27) of MT neurons and 6% (4/33) of MST neurons. A 50% increase in the response was observed in 41% (11/27) of MT neurons and 27% (9/33) MST neurons.

Given that we studied the responses of the same neurons on long-duration trials as well as short-duration trials in *monkey N*, we were able to directly compare the modulation due to attention in these same neurons. Because the attention index is a highly nonlinear transformation of the responses, to make this comparison we calculated the difference in the firing rate between the attended and unattended trials on the short- and long-duration trials for the same neurons shown in Fig. 3. Figure 4 plots this difference in activity for the short-duration trials ( $x$ -axis) against the long-duration trials ( $y$ -axis) for neurons recorded in MT (●) and MST (■). The thin line shows perfect correlation between these two measures, whereas the heavy line shows the best-fit regression line for the data pooled between both cortical areas. Most points fell above the line of unity slope, indicating that the majority of neurons showed a greater difference in the activity on the long-duration trials as

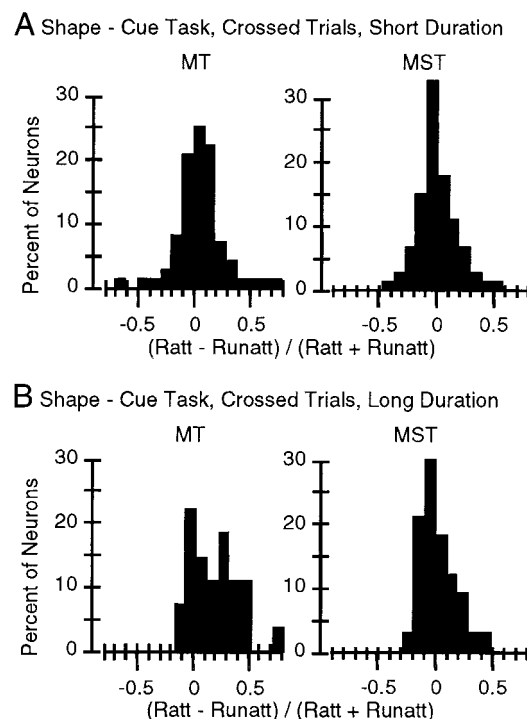


FIG. 3. Frequency distribution of the attention index during the crossed trials of the shape-cue task. *A*: attention indexes measured across the population of middle temporal (MT) neurons (*left*) and medial superior temporal (MST) neurons (*right*) during the short-duration trials. Population indexes were not significantly different between monkeys, and therefore the data were pooled. These frequency distributions were not significantly different from a population with a mean of zero (mean  $\pm$  SD:  $0.03 \pm 0.21$  and  $-0.01 \pm 0.21$  for MT and MST, respectively). *B*: attention indexes for MT (*left*) and MST (*right*) neurons during the long-duration trials. For both cortical areas, only the long-duration trials showed a consistent effect depending on the direction the target stimulus was moving ( $0.17 \pm 0.23$  and  $0.15 \pm 0.16$  for MT and MST, respectively). Ratt, mean response when the target was moving in the best direction in the RF of the neuron and the other stimulus was moving in the null direction; Runatt, mean response when the target was moving in the null direction and the other stimulus was moving in the best direction.

compared with the short-duration trials. This analysis shows that the difference in the population distributions of Fig. 3 reflect a consistent difference in the change in activity on a neuron-by-neuron basis depending on the trial conditions.

In summary, these results showed that there were significant differences in neuronal responses between the attended and unattended stimuli in our paradigm only when we used the long spatial/temporal stimulus configuration. We did not find substantial differences in the effect of attention between MT and MST.

#### Attentional effects on directional tuning

So far we have only considered the effect of attention on the response of the neuron to motion in the best direction, but we also determined the effect of attention on the relative strength of the response in the seven other nonbest directions tested. For this analysis we calculated the vector strength on the attended and unattended trials and compared these vector strengths using a contrast ratio (see METHODS). The results of this analysis for the same neurons described in Fig. 3 are shown in Fig. 5. For the short-duration trials (Fig. 5A) in both MT and MST the population distribution was not statistically significantly dif-

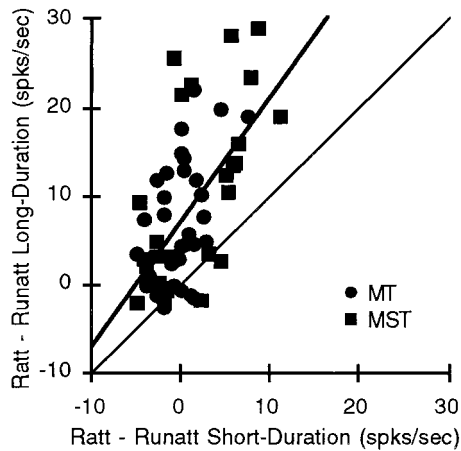


FIG. 4. Comparison of the attention effect between short- and long-duration trials. The difference in activity between the attended and unattended trials is shown for each of the MT (●) and MST (■) neurons studied using both the short-duration (x-axis) and long-duration (y-axis) trials. The thin line is drawn through unity slope. The thicker line shows the best fit regression line (slope = 1.39; intercept = 7.0;  $r = 0.594$ ). The majority of neurons showed a larger difference in activity for the long-duration trials compared with the short-duration trials.

ferent from a population with a mean of zero, which was expected due to the lack of any attention effect for the vast majority of these neurons. For the long-duration trials, in which an effect of attention was observed, the population distribution again was not statistically significantly different from a population with a mean of zero. This indicates that the effect of attention under these experimental conditions is an elevation of the response for all directions of motion, and is not selective for the best direction of motion.

#### Paradigm parameters alter the modulation

A number of previous experiments, beginning with those in extrastriate area V4 (Moran and Desimone 1985), have found that the influence of attention is best observed when two stimuli are presented in the RF of the neuron at the same time, as in our crossed trials. To address this issue in the present pursuit-based experiments, we compared the neuronal response on attended and unattended trials when one stimulus passed through the RF and the other was in the opposite hemifield. These uncrossed trials (Fig. 1C) were interleaved with the crossed trials considered above so that the analysis could be performed on the same neurons. Figure 6 shows the response on uncrossed trials for the same neuron shown in Fig. 2. For both the short-duration trials (Fig. 6A) and the long-duration trials (Fig. 6B), there was very little apparent difference in the activity of the neuron whether the stimulus moving through the RF of the neuron was attended to or not. As expected, there was no statistically significant difference in the responses between the attended and unattended trials for either duration stimulus (unpaired  $t$ -test;  $P > 0.05$  for both short- and long-duration trials). Across the population of studied neurons, in area MT there were 10/72 neurons (14%) with statistically significant differences in activity, with 7 showing decreases and 3 showing increases with attention on the short-duration trials. For long-duration trials, 5/27 neurons (19%) showed a difference, with 4 neurons having increased activity and 1 neuron having decreased activity with attention. For MST

neurons, on the short-duration trials 7/73 neurons (9%) showed a statistically significant difference, with 4 showing increases and 3 decreases in activity with attention. On the long-duration trials, only 3/33 (10%) showed any difference in activity, and these neurons all showed increases with attention. Figure 7 shows the attention index for both areas MT and MST. None of these sample distributions were significantly different from a population with a mean of zero ( $t$ -test;  $P > 0.05$ ), and these distributions were also not statistically significantly different from each other ( $P > 0.05$ ). Thus the greatest effect of attention that we observed in the shape-cue task occurred when the selection was between two stimuli crossing the RF of the neuron, and there was a relatively long spatial/temporal interval before the stimuli cross the RF. The effect was equal in MT and MST.

In this analysis of the uncrossed trials (Fig. 7) as well as that for the crossed trials (Fig. 3) we compared the total activity during the period from the onset of stimulus motion to the start of the eye movement. If these neurons had a consistent attentional effect with a long latency, there should be a divergence of the two plots following that latency (e.g., see Motter 1994a; Seidemann and Newsome 1999). Alternatively, if the attentional effect was only in the early phase of the response, the two plots should initially be distinct, and then become similar

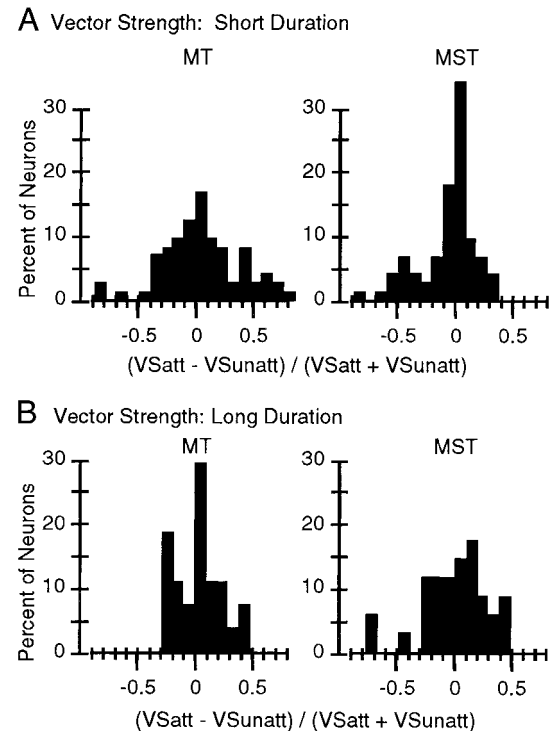


FIG. 5. Effect of attention on the directional selectivity of neurons. *A*: vector strength tuning measurements during the short-duration crossed trials for the population of MT neurons (*left*) and MST neurons (*right*) shown in Fig. 3. VSatt, vector strength when the target was moving in the best direction and the other stimulus was moving in each of the 7 nonbest directions; VSunatt, vector strength when the target was moving in 1 of 7 nonbest directions and the other stimulus was moving in the best direction. These populations were not significantly different from a population with a mean of zero (mean  $\pm$  SD:  $0.04 \pm 0.35$  and  $-0.01 \pm 0.26$  for MT and MST, respectively). *B*: vector strength tuning measurements on the long-duration trials for MT neurons (*left*) and MST neurons (*right*). Conventions as in *A* (mean  $\pm$  SD:  $0.05 \pm 0.24$  and  $0.01 \pm 0.28$  for MT and MST, respectively). The directional tuning did not change depending on which stimulus was the target.

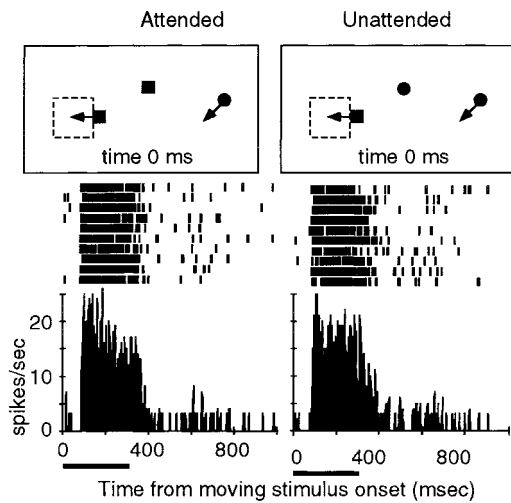
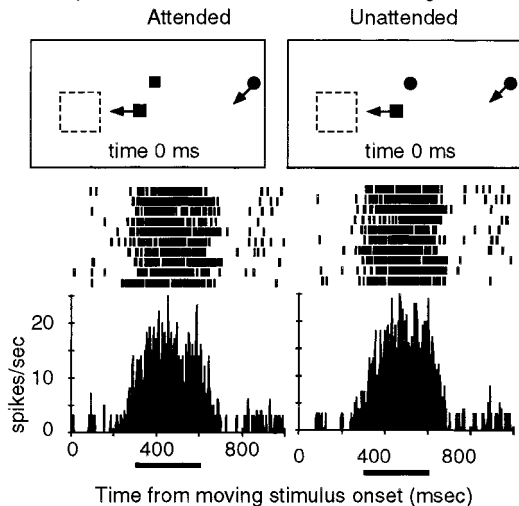
**A** Shape - Cue Task, Uncrossed Trials, Short Duration**B** Shape - Cue Task, Uncrossed Trials, Long-Duration

FIG. 6. Example responses from the same single MT neuron shown in Fig. 2 during the interleaved uncrossed trials of the shape-cue task. *A*: responses during the short-duration trials. Rasters (*top*) and PSTHs (*bottom*) when the target was moving in the best direction of the neuron (*left*) and when the same stimulus was presented in the receptive field but the target was moving in the opposite visual hemifield (*right*). Conventions as in Fig. 2. *B*: response during the interleaved long duration trials. Rasters (*top*) and PSTHs (*bottom*) from the same neuron when the moving stimuli initiated their motion at a greater spatial separation (see Fig. 1). There was no difference in the neuronal activity depending on which stimulus is the target for the eye movement. This neuron had an attention index of  $-0.01$  for the short-duration and  $-0.02$  for the long-duration trials.

after the period of the attentional effect was over. To determine whether there was a consistent change in the level of activity throughout the trial, we pooled the responses of the population of neurons. Such a pooling procedure is shown in Fig. 8 for all MT neurons that showed a statistically significant modulation. For the short-duration crossed trials (Fig. 8A), the population activity was essentially identical throughout both the attended and unattended trials. In contrast, for the long-duration crossed trials there was an elevation of activity that began close to the beginning of the response, reached a plateau, and then decreased throughout the rest of the trial, largely in parallel with the response to the unattended stimulus. For the uncrossed

trials (Fig. 8B), both the short-duration and long-duration conditions showed virtually identical responses for the attended and unattended stimuli. Importantly, there was no clear difference in the responses as a function of time from the onset of the stimulus, indicating that the differences in the activity of these neurons were modulated consistently across the entire stimulus period. Similar analysis of MST neurons showed the same result. Pooling the responses across neurons that did not show a statistically significant effect did not show a difference in the response at any period during the trial under any stimulus conditions. We also analyzed these responses using 50- and 100-ms windows throughout the response period and found essentially the same result (data not shown). Therefore we did not observe an increased response limited to one period of the visual response, but rather an increased response on attended trials that followed the same time course as that on the unattended trials.

Any movement of the eyes to the target would not have affected the response in these experiments because we terminated the period in which the visual response was analyzed before the onset of the saccade. We also considered the possibility that the difference in activity could be due to small differences in the eye position before the saccade, which would result in a different retinal image depending on which stimulus was the target. These monkeys were required to maintain

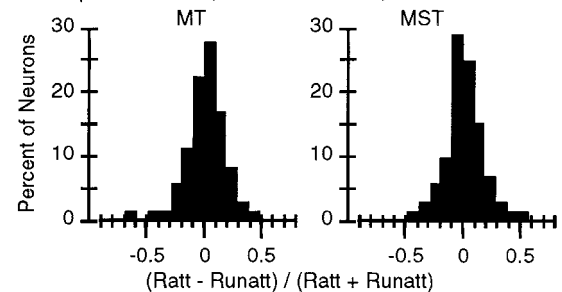
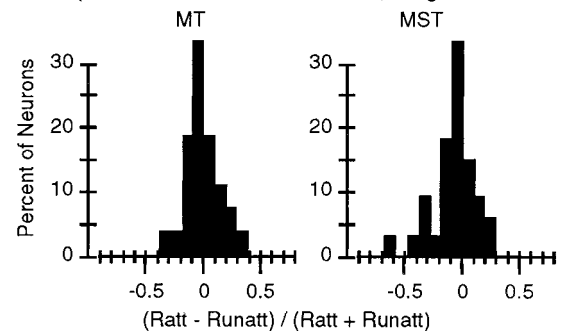
**A** Shape - Cue Task, Uncrossed Trials, Short-Duration**B** Shape - Cue Task, Uncrossed Trials, Long-Duration

FIG. 7. Frequency distributions of the attention index calculated for the shape-cue task uncrossed trials. *A*: attention index for the short-duration trials for MT (*left*) and MST (*right*) neurons. *B*: attention index measured for the long-duration trials for MT (*left*) and MST (*right*). Data shown are only from trials where the stimulus was moving through the receptive field of the neuron in the best direction. There was no significant difference in the activity of neurons across the population depending on whether the stimulus in the receptive field was the target for the eye movement or not (mean  $\pm$  SD. *A*:  $-0.01 \pm 0.21$  and  $0.02 \pm 0.19$  for MT and MST, respectively; *B*:  $-0.04 \pm 0.20$  and  $-0.02 \pm 0.20$  for MT and MST, respectively). Ratt, mean response when the target was moving in the best direction in the RF of the neuron; Runatt, mean response when the target was moving in the opposite visual hemifield and the 2nd stimulus was moving in the best direction in the RF of the neuron.

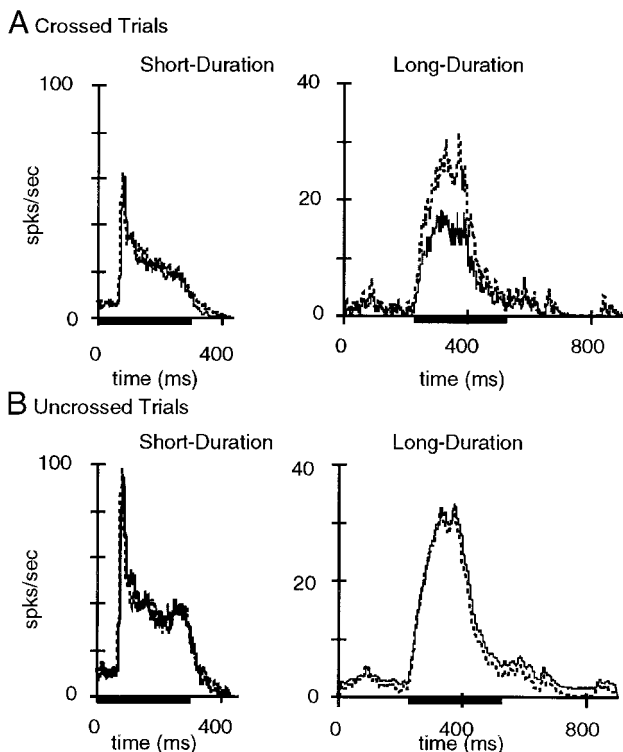


FIG. 8. Responses of the population of neurons during the attended and unattended conditions. *A*: responses during the crossed trials of all neurons showing a statistically significant difference (see RESULTS) throughout the time period of analysis. Heavy bar on the *x*-axis shows the approximate time that the stimulus was in the receptive field. The attended response (heavy line) and unattended response (thin line) were essentially superimposed on the short-duration trials (*left*) but not on the long-duration trials (*right*). *B*: same analysis for the uncrossed trials. On these trials the responses were essentially superimposed on both the short-duration (*left*) and long-duration (*right*) trials.

fixation throughout the data collection period; however, the eye position criteria was set at  $\pm 2^\circ$  due to the large size of the fixation stimulus. We therefore compared the eye position measured at the start of the trial (*time 0* ms of Fig. 1) and the eye position measured before the saccade on the attended and unattended trials. We restricted this analysis to those trials in which the target moved in either the best or null direction, because these opposite directions should yield the greatest difference if the monkey initiated an eye movement before the saccade. We found no statistically significant difference between attended and unattended trials for either time period when measured across the population, or when restricted to only those neurons that showed a statistically significant effect of attention (*t*-test,  $P > 0.05$ ). This finding, coupled with the time course analysis of Fig. 8 indicates that small eye movements during the fixation period cannot account for the attentional effect we observed.

Another variable in our experiment is the type of cue used to indicate which stimulus was to be the target for the eye movement and when it was given. In the shape-cue task, the monkey was cued which target shape was to be used on that trial but not where that target was to be in the visual field. In a separate set of experiments, we used a location cue rather than a shape cue. The location cue task (Fig. 1*D*) addressed the possibility that shifting attention to the appropriate region of visual space before the onset of the moving stimuli would generate a difference in the neuronal response. In this task the

monkey was cued which hemifield the target stimulus would appear in, similar to other recent experiments (Seidemann and Newsome 1999; Treue and Maunsell 1996, 1999). We were only able to use uncrossed trials in this paradigm, because the direction of motion of the target stimulus was not cued, only the hemifield location. Figure 9 shows the responses of the same MT neuron illustrated in Figs. 2 and 6 during the location cue task in which the cue was extinguished once the monkey fixated, and the moving stimuli then appeared after a 300- to 600-ms fixation period. The neuronal responses shown in Fig. 9 indicate that there was very little difference in the response of this neuron even when the attention was directed to the RF location by an earlier location cue (2-tailed unpaired *t*-test;  $P > 0.05$  for both short- and long-duration trials). This lack of attentional modulation was consistently seen across the sample of both MT and MST neurons in both monkeys, as well as for

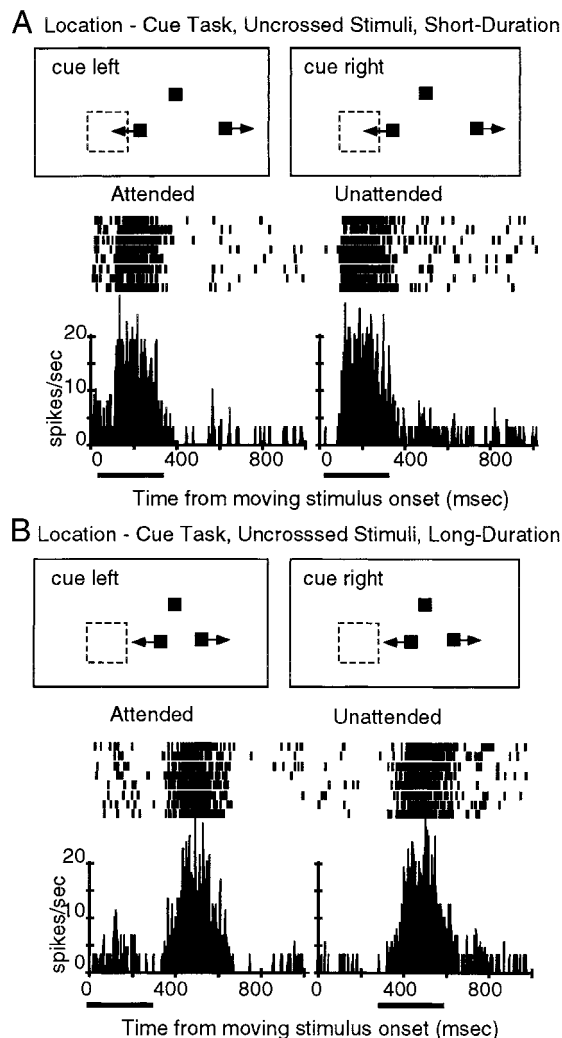


FIG. 9. Single neuron response during the location-cue task. Data taken from a different session from the same neuron illustrated in Figs. 2 and 6. *A*: response from the neuron during the short-duration trials where the stimulus moving in the best direction was the target (*left*) and when the stimulus moving in the opposite visual field was the target (*right*). *B*: response from the same neuron to the same stimulus configuration on the long duration trials. Conventions as in Fig. 2. The attention index for these trial types in this cell were  $-0.04$  for the short-duration trials, and  $-0.01$  for the long-duration trials. There was no consistent difference in the activity of this neuron during the stimulus period depending on which stimulus was the target.



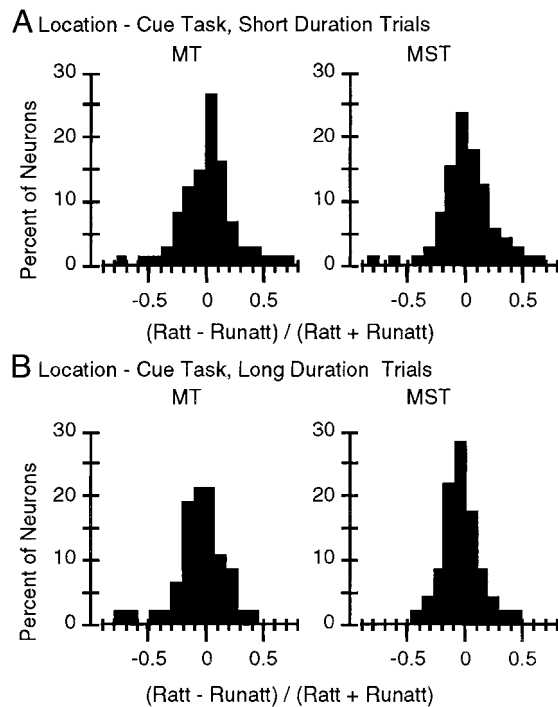


FIG. 10. Frequency distribution of the attention index measured during the location-cue task. *A*: attention indexes measured from MT neurons (*left*) and MST neurons (*right*) during the short-duration trials. *B*: attention indexes measured from MT (*left*) and MST (*right*) neurons during the long-duration trials. None of these population responses was statistically significantly different from a population with a mean of zero (mean  $\pm$  SD, *A*:  $0.02 \pm 0.19$  and  $0.00 \pm 0.24$  for MT and MST, respectively; *B*:  $0.04 \pm 0.24$  and  $0.07 \pm 0.19$  for MT and MST, respectively).

both the short-duration and long-duration versions of this task. For MT neurons, only 6/75 neurons (8%) showed any modulation on the short-duration trials (5 increases and 1 decrease) and only 3/47 (6%) on the long-duration trials (2 increases and 1 decrease). In MST, only 5/72 neurons (7%) showed modulation on the short-duration trials (2 increases and 3 decreases) and only 4/45 (9%) showed modulation on the long-duration trials (2 increases and 2 decreases). The frequency distribution of the attention index emphasizes this result, and under these stimulus conditions the population of MT or MST neurons was not statistically significantly different from zero (Fig. 10). Vector strength analysis as described above for the short- and long-duration shape-cue task again showed that the population of MT and MST neurons was not statistically significantly different from a population with a mean of zero ( $P > 0.05$ ). Finally, analysis of the differences in activity as a function of time from stimulus onset showed the same result as for the uncrossed trials of the shape-cue task, namely, the lack of any consistent difference in the responses throughout the stimulus period.

In summary, these results indicate that attending to the visual field location corresponding to the RF of the neuron under study before the onset of the stimulus motion did not significantly effect the firing rate of the vast majority of neurons and indicates that the modulation observed on the long-duration crossed trials of the shape-cue task cannot be simply explained by a shift of attention to the visual field containing the target stimulus. The results support the hypothesis that the long spatial/temporal delay for the stimulus to reach the RF is

not in itself sufficient to produce the observed modulation but that it probably must be paired with multiple stimuli in the RF.

#### Contribution of changes in spontaneous activity

Experiments on the effects of attention in area V4 and inferior temporal cortex (Luck et al. 1997) have indicated that attention to the stimuli can alter the spontaneous activity of the neurons, and have suggested that this change in activity is indicative of a change in the excitability of the neurons during the attention task. We also saw such changes in spontaneous activity during some experiments, and we therefore compared the spontaneous activity just before the onset of the visual stimuli when attention was directed toward and away from the RF location across our sample of neurons. We compared the firing rate over the 150 ms before the onset of the moving stimuli on attended trials to the same period before unattended trials. For the shape cue task, only trials with a 300- to 500-ms fixation period were used, which was the fixation period used during the location-cue task, and inspection of the postcue period before the onset of the moving stimulus showed no indication of an offset or other stimulus-related response during this 150-ms period. We found no difference in spontaneous activity for either the short- or long-duration trials during the shape-cue task, which is reassuring given that the monkey had no knowledge of the hemifield in which the target stimulus would later appear. In contrast, for the location cue trials, we found a small but statistically significant increase in the spontaneous activity on trials in which the monkey was cued that the target would move through the RF of the neuron under study. This increase varied between neurons, and the number of trials provided a strong statistical sample (256–320 trials/neuron), but in the vast majority of cases this increase in activity was quite small. For MT neurons, the population average went from  $3.47 \pm 2.97$  (SD) spikes to  $4.45 \pm 3.75$  spikes, with 37/75 individual neurons showing significant differences (all increases) during trials in which the target was to appear in the RF of the neuron (1-tailed unpaired *t*-test,  $P < 0.05$ ). There was a similar finding for MST neurons, where the population average increased from  $2.78 \pm 1.91$  spikes to  $3.55 \pm 2.61$  spikes, with 34/72 individual neurons showing a statistically significant difference (again, all increases) during trials in which the target was to appear in the RF of the neuron.

To show these differences across our sample of neurons, we used the same attention index for the 150-ms prestimulus period for both MT and MST neurons. As expected, the population of neurons did show an average attention index greater than zero (Fig. 11). For area MT neurons, the population average was  $0.14 \pm 0.22$  and  $0.07 \pm 0.14$  for *monkeys N* and *P*, respectively. For area MST neurons, the population average was  $0.05 \pm 0.24$  and  $0.19 \pm 0.27$  for *monkeys N* and *P*, respectively. Overall, however, these differences in spontaneous activity were small and were not statistically significantly correlated with the difference in the driven activity when measured across neurons in our sample ( $r = 0.03$ ;  $P > 0.05$ ).

#### DISCUSSION

We observed attentional modulation of the responses of MT and MST neurons to stimulus motion, but only on trials in which the two stimuli both moved through the RF of the

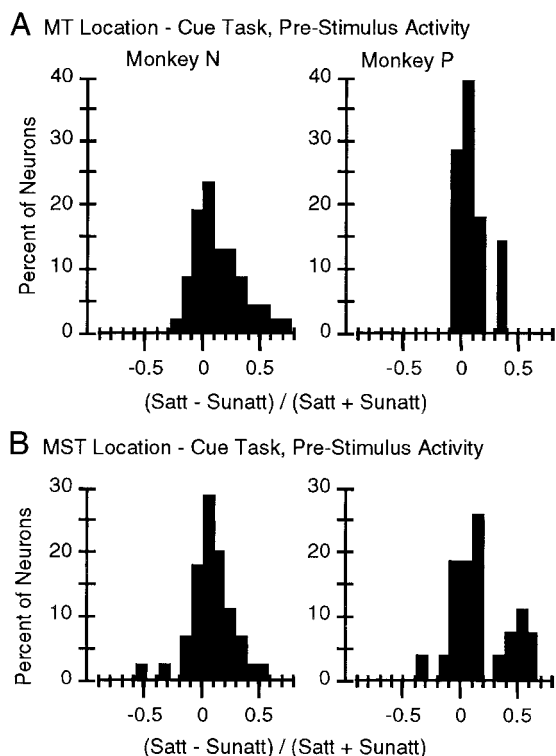


FIG. 11. Frequency distribution of the attention index of the spontaneous activity measured during the location-cue trials. Because both long- and short-duration trials were randomly interleaved and the monkey could not predict which trial type was about to occur, both trial types were pooled. Population responses were different between the 2 monkeys and are therefore shown separately. *A*: MT neurons. *B*: MST neurons. These populations were statistically significantly different from a population with a mean of zero (mean  $\pm$  SD, *A*:  $0.14 \pm 0.22$  and  $0.07 \pm 0.14$  for monkeys *N* and *P*, respectively; *B*:  $0.05 \pm 0.24$  and  $0.19 \pm 0.27$  for monkeys *N* and *P*, respectively). Satt, spontaneous activity before trials where the target was the stimulus that would move through the receptive field; Sunatt, spontaneous activity before trials where the target was not the stimulus that moved through the receptive field.

neuron, and only when these stimuli moved for a longer time and distance before the pursuit initiation. Attention acted on all directions of motion so that the directional tuning curve remained the same. The modulation was substantially the same in MT and MST. We will first consider the relationship between the current study and our previous study (Recanzone and Wurtz 1999), and then how our observations relate to models of competitive interactions by attention, consider why our attentional effects were relatively small, and finally compare a few of our salient observations with those in previous attentional experiments in extrastriate cortex.

#### Relationship to our previous study of the population models

The present experiments continue the analysis of the activity of MT and MST neurons during pursuit initiation begun in the previous report (Recanzone and Wurtz 1999). In those experiments, we found that both the monkey's pursuit behavior and the correlated neuronal activity could be described by either a vector average of the two directions of motion or a winner-take-all in which the response to one direction prevailed depending on the stimulus conditions. The winner-take-all model (for both behavior and neuronal modulation) was the better

predictor on trials with a longer spatial/temporal interval before the two stimuli crossed the RF of the neuron, whereas the vector average model was the better predictor for the shorter spatial/temporal interval trials. One concern raised by those experiments was that an attention effect could account for the results, and indeed recent studies have shown striking modulation of MT and MST neurons as a function of attention (Treue and Maunsell 1996, 1999), although other studies have shown limited modulation (e.g., Ferrera and Lisberger 1997; Newsome et al. 1988; Seidemann and Newsome 1999). In the present experiments we addressed the issue of attention directly by analyzing the MT and MST neuronal responses when the monkey did and did not attend to a moving stimulus. The effect of attention is best addressed by comparing the results of our previous experiments with the effect of attention in the present experiments as measured by the attention index.

The relationship between the vector average prediction and the attention index is relatively straightforward, because the vector average model predicts the same response under both the attended and unattended conditions. The attention index we used (the difference in the average response between the attended and unattended trials divided by the sum of the 2 responses) gives values varying from  $-1.0$  (no attended response) to  $+1.0$  (no unattended response), with zero being no difference in the response between the two conditions. When the vector average model better predicted the neuronal response (the short-duration crossed trials), the responses should have an attention index near zero, which is what we observed.

The only time we saw modulations that could be related to the monkey's attention to a stimulus was when the winner-take-all model better predicted the behavior and neuronal activity. The relationship between the winner-take-all prediction and the attention index, however, is less straightforward because the attention index is a nonlinear transformation of the data. An attention index of 1.0 will only occur when there is no response during the unattended trials, which was not observed in either our uncrossed or our crossed stimulus condition. For example, even for neurons that respond with 100 spikes for stimuli moving in the best direction, and 10 spikes for stimuli moving in the null direction, a winner-take-all model predicts a response of 100 spikes with an attention index of only 0.82 (90/110). Therefore the winner-take-all model still predicts some response during the unattended trials, and given the relatively high firing rates for the crossed trial stimuli (e.g., Fig. 2) (Recanzone and Wurtz 1999) an attention index greater than zero, but still  $<1$ , is entirely consistent with the winner-take-all model. It is also difficult to compare directly across our two studies because the previous study used the *uncrossed* trials to predict the responses on the *crossed* trials, whereas in the present study we compared the responses of the neurons with the identical visual stimulus moving through the receptive field, so that in the two studies different trial types were compared. Nonetheless, the results of the present experiment are consistent with those using uncrossed trials to predict the response of the crossed trials. For the long-duration crossed trials, 48 and 36% of the MT and MST neurons, respectively, showed statistically significant differences in activity consistent with a winner-take-all model. For neurons tested under both the short- and long-duration conditions, there was a consistently greater difference in activity on the crossed long-duration attended trials compared with the short-duration trials,

indicating that this difference in the attentional effect was consistent across neurons.

Taken together, these results show that the trial conditions (long-duration crossed trials) in which there is the strongest tendency for the winner-take-all model to better describe the activity of MT and MST neurons are the same conditions under which the effect of attention is most prominent. It is clear that an attentional mechanism alone cannot account for these results given the lack of attentional modulation observed under other trial conditions: the short-duration trials with the fixation stimulus as the cue, the long-duration trials in which only one stimulus moved through the receptive field of the neuron, or the location-cue trials in which the monkey's attention was to the region of visual space corresponding to the RF of the neuron. The attention effect may be necessary for the shift from an average to a winner-take-all mode, but it is not sufficient. On the other hand, it seems clear that the modulation of neuronal activity, which we have related to attention, is not related to the preparation to move the eye because the modulation varied across conditions even though the eye movement always occurred.

Probably the simplest way to view the results of the present and the previous experiments is in their contribution to the winner take all response mode. In the previous experiments, the longer time of motion before pursuit was enough to allow one stimulus to produce a stronger response in the neuron under study. In the present experiments, the effect of attention to one stimulus was to strengthen the neuronal response to that stimulus. Thus both effects acted to strengthen the response to the one stimulus thereby contributing to the winner take all mode of the response, and the underlying mechanisms are likely to be similar in both cases.

#### *Competitive models of attention*

The effects of attention in MT and MST in our experiments are consistent with neuronal competition models of attention and possibly also place limits on these models. In the biased competition model of Desimone and Duncan (1995), when two stimuli appear, they activate two populations of neurons that compete with each other. Which neurons are active is determined by the effectiveness of the stimuli activating them (a bottom-up effect). The effect of attention is to bias this competition toward one population or the other (a top-down effect). These two populations must be close together for lateral excitatory or inhibitory interactions to be effective, and this aspect of the hypothesis arises from the observation that for the attention to be most effective on the responses of cortical neurons, the two stimuli must both fall within the RF of the neurons [such as in the original V4 and IT observations of Moran and Desimone (1985)]. In our experiments, we found the greatest effect of attention when both the attended stimulus and the other stimulus fell in the RF of the neuron (the crossed trial condition). Our results are also consistent with a series of experiments showing larger attention effects with two stimuli in the RF in other extrastriate areas (Luck et al. 1997; Moran and Desimone 1985; Motter 1994a,b; Reynolds et al. 1999; Treue and Martinez Trujillo 1999; Treue and Maunsell 1996, 1999) but not those of Seidemann and Newsome (1999).

We think our experiments also add a constraint on the characteristics of the competing neuronal populations: just as

the competing populations cannot be too far separated, they also cannot be too close together. The results of the present study show that when the two populations are initially very close together in both stimulus space and time, in our case with similar receptive field locations but different best directions, attention is much less effective in acting on one population. We saw the largest attention effect when we also saw a winner-take-all mode indicating that the neuronal activity of two subpopulations of activity could be separated, but when the two populations were so close together that the output was simply an average, attention was not effective in acting on a subpopulation. The populations must therefore have some degree of separation that allows competition between them and then attention can introduce a bias. Although we have seen this effect in the motion processing neurons in MT and MST, a similar principle of separation of subpopulations in order for attention to act might be present in the other extrastriate areas where attention for stimulus form has been identified and in other areas more closely related to saccade generation such as posterior parietal cortex (e.g., Bushnell et al. 1981; Lynch et al. 1977; Robinson et al. 1995; Steinmetz et al. 1994) and superior colliculus (Goldberg and Wurtz 1972; Kustov and Robinson 1996).

Thus the populations on which the competition acts must be close enough to be in one RF but not so close that the two populations are indistinguishable. We previously interpreted the winner-take-all behavior to reflect hypothetical excitatory intracortical connections between columns of neurons with the same best direction and with RFs along the trajectory of the stimulus, and the vector average behavior to reflect the interactions between neurons with similar RF locations but different best directions (see Recanzone and Wurtz 1999, Fig. 13). This organization would be part of the bottom-up processing within MT and MST on which the top-down influences of attention would act.

#### *Attentional modulation in MT and MST*

Our current study was as striking for the number of conditions in which we did not see modulation as those in which we did, and even when we did see modulation, the effects were small compared with those observed in MT and MST for a speed change detection task (Treue and Maunsell 1996). In our experiments, the median increase in response in the long-duration crossed trials was 11% for MT and 16% for MST, whereas for the best case in the Treue and Maunsell experiments the median modulation was 86% for MT and 113% for MST. For a direction of motion discrimination task in MT, Seidemann and Newsome (1999) found even less modulation than we did, a median of 8.7%. The magnitude of attentional modulation is clearly dependent on the conditions under which it is invoked and, although it is difficult to compare across very different experimental paradigms, it is worth noting the similarities and differences that we think are important along with a few that we think probably are not.

Comparing recent results of attentional effects in visual motion processing areas indicates that at least two factors appear to be critical for producing attentional modulation in MT and MST. One factor is the nature of the cue, which in our experiments was either the shape of the fixation stimulus or the location in the visual field that the target would appear in, but

in neither case was the direction of the target motion indicated before the trial began. We observed the largest attentional effect on the long-duration crossed trials of the shape-cue task, where the monkey presumably identified the target and its direction of motion before it entering the RF of the recorded neuron. The larger attention effects in the Treue and Maunsell (1996, 1999) experiments were observed under cueing conditions that provided both the location and the impending direction of motion of the target. In their task, the stationary cue was offset from the center of the RF and began its motion to the nearest edge of the RF before reversing direction, and given the unvarying trajectory of the stimuli throughout each experimental session, the monkey likely predicted both the starting direction of motion and when the target stimulus was about to change direction. The small effects in the direction discrimination task of Seidemann and Newsome (1999) were observed when only cueing the location of the stimulus to be discriminated. Thus a key difference is that the Treue and Maunsell design would allow an attentional mechanism to engage the specific population of neurons that would provide the most salient information about the stimulus (those with the appropriate best direction) before the start of the trial, whereas in our experiments which neurons would be providing this information remained unknown until after the shape discrimination was made. Furthermore, if it takes time for the top-down attention effect to develop, our experiments would have minimized the period over which this attention could act. This is further supported by recent experiments by Treue and Martinez Trujillo (1999), who similarly used a location cue task and saw significant modulation of most of the MT neurons encountered, but they analyzed the responses in a period of 200–1,200 ms after the onset of target motion. Other studies that found little modulation in MT also indicated the stimulus to be attended for only a short time before the movement or discrimination was required (Ferrera and Lisberger 1995, 1997; Newsome et al. 1988). At least in these motion-based tasks, a location cue alone might be an inadequate one because it does not specify the population of neurons related to the relevant stimulus parameter, in this case the direction of motion.

The other factor is whether both stimuli were in the RF of the neuron. In our experiments we only saw a modulation when both stimuli passed through the RF, and in the Treue and Maunsell (1996, 1999) experiments the effect was substantially stronger when two stimuli were within the RF. These findings support the competition model of attention described above, in that there must be two competing populations of neurons activated by the stimuli before these top-down influences can be observed. This is confounded somewhat by the relatively small modulations seen with both one and two stimuli in the RF of the neuron observed by Seidemann and Newsome (1999). The interactions of the type of cue (in this case the direction of the target vs. the target location) and the stimulus configurations may account for these differences, although clearly more experiments are necessary to resolve this issue.

A final consideration of why the effects we observed were smaller than those of Treue and Maunsell (1996, 1999) might be the extent to which attention was concentrated on one stimulus. In the task that we used most, the shape-cue task, the monkey had to quickly perform a shape discrimination of two moving visual stimuli, and was cued to initiate pursuit by the offset of the fixation stimulus. These circumstances are thus

quite different from those in which the monkey was required to attend to only a particular visual stimulus, and make a lever press response (Treue and Maunsell 1996). Our monkeys had to attend to the moving stimuli as well as the fixation stimulus, and this division of attention may have contributed to the more limited modulation we saw.

Finally there are two issues that we do not think are critical in producing the attentional modulation. The first is whether our use of a movement rather than a discrimination could account for our observations of relatively limited modulation. We consider this very unlikely because all of our tasks required a movement, but attentional modulation was associated with only a few conditions. It is also worth noting that because a number of our paradigms produce little attentional modulation, whereas all of them produced pursuit movements, preparation to move seems to be a less likely correlate of the change in activity than does the effect of attention. In addition, both the Treue and Maunsell (1996) and Seidemann and Newsome (1999) experiments required discriminations rather than motion-guided movement, but one produced much stronger and the other weaker modulation than we observed.

The second issue is related to the difficulty of the task: if the task were easy, no attention might be required. In our experiments the monkeys were required not just to detect some change in the visual motion signal but to exactly match its eye speed and direction to that motion. This is arguably a more demanding task than just detecting a change in the motion, and the effect of attention ought to be commensurate with this demand. This is clearly not a determining factor because the tasks in which we observed an attention effect were nearly identical in difficulty to those in which we did not, although difficulty might influence the strength of the attention effect.

In summary, we think that the effectiveness of the cue in engaging attention, combined with a competitive model of attention, can account for the results of this and the preceding experiment (Recanzone and Wurtz 1999). On the long-duration trials the attentional effect was most prominent and the winner-take-all model was a better predictor of the response because attention had an opportunity to be engaged. In contrast, we saw little effect of attention and neuronal responses better predicted by the vector average model on the short-duration trials. On the uncrossed trials, we saw no attentional effect (using either the shape or location cue) presumably because the two competing populations were too far apart for attention to influence one population over the other. If our reasoning is correct, cueing the monkeys with both the location and direction of motion of the target before the onset of the trial should result in a greater attention effect and both neuronal responses and eye movement metrics that are even more closely predicted by the winner-take-all model than were described under the present experimental conditions (Recanzone and Wurtz 1999).

#### *Comparison to other studies of attention in extrastriate cortex*

Attentional modulation has been observed throughout visual cortex in the ventral pathway (V1, V2, V4, and IT) (e.g., Haenny and Schiller 1988; Luck et al. 1997; McAdams and Maunsell 1999; Moran and Desimone 1985; Motter 1993, 1994a,b; Reynolds et al. 1999; Richmond and Sato 1987), and the dorsal pathway (MT and MST) (e.g., Ferrera and Lisberger

1997; Newsome et al. 1988; Seidemann and Newsome 1999; Treue and Maunsell 1996) including the parietal cortex (Bushnell et al. 1981; Lynch et al. 1977; Robinson et al. 1995; Steinmetz et al. 1994). Several observations in the present experiments are worth noting with respect to these other attention studies in extrastriate cortex.

The first is the specificity of the attention effect. Studies of attention in area V4 have shown that attention alters not just the response to stimuli with the best orientation for the neuron, but the response to the nonbest orientations as well (McAdams and Maunsell 1999; see also Motter 1993). The comparable measure for the neurons sensitive to motion in MT and MST would be the direction tuning curve, and we found that attention had little influence on the directional tuning of both MT and MST neurons. This result is in agreement with Treue and Martinez Trujillo (1999), who tested MT neurons using random dot patterns as the visual stimulus. Thus for both orientation and for direction, attention altered the activity not of just the best stimulus but the range of stimuli to which the neuron responded and thus did not change the tuning of the neuron studied. In both cases attention acts to enhance the information carried by the neuron, not to alter its content.

The second is the relation of changes in the background activity of the neuron to the magnitude of the attention effect. Luck et al. (1997) found that larger attention effects in area V4 were accompanied by increases of background activity, and they interpreted these changes as an indicator of the changes in the underlying excitability of the neurons. We also found a significant change in the background discharge rate but found that the changes were small and unlikely to be sufficient to explain the attentional modulation. Our results would suggest that these are parallel events rather than one being an indicator of the other. But it is also possible that weaker attention effects that we observed in MT and MST are accompanied by a lower level of background activity.

Perhaps the most striking finding, and one that we have considered above, is the inference that attention takes time to develop. The clearest case of this is the minor attentional effect in the short-duration shape-cue task when only 150 ms intervened between the identification of the object of attention and the movement and the more robust attentional effect in the long-duration shape-cue task when the intervening time was 450 ms. On the short-duration trials, there was no modulation 300 ms after stimulus motion onset, and by that time the stimuli were leaving the RF. Such a latency for the attention effect to develop was also clearly observed by Seidemann and Newsome (1999), who saw the modulation at ~300 ms after stimulus onset, similar to observations in V4 (Motter 1994a,b).

The latency of the attention effect could also interact with the size of the RF of the neurons when motion stimuli are used. This factor could potentially account for the roughly equivalent modulation of MT and MST neurons in contrast to the larger effect on MST than on MT neurons described by Treue and Maunsell (1996) and the lack of modulation in MT seen by Newsome et al. (1988). This would occur if attentional effects in MST had a similar time course to those in MT but the stimuli started at the edge of the RF in both areas; the stimuli would be out of the RF for those neurons with small RFs sooner than those with larger RFs. If we had allowed a longer time between target identification and movement initiation, we might have

seen more substantial attentional modulation in MST than in MT because the added time would have allowed further motion within the RF. Thus we think that a significant factor in the strength of attentional modulation is the time taken for the top-down attention effect to develop, and that this time should be measured not just from stimulus onset but from the cue for attention.

Thus it seems reasonable to conclude that just as the bottom-up activity has a visual latency that is dependent on the stimulus conditions, the top-down attention modulation also has a latency that is dependent on task conditions. The attention latency may usually be longer than the visual latency, but can be engaged before the appearance of a visual stimulus if the salient features of the target can be predicted. The attention latency must be measured from the onset of the cue rather than the onset of the stimulus, and the nature of that cue undoubtedly alters the attentional latency.

We thank our colleagues at the Center for Neuroscience (particularly K. H. Britten, H. Heuer, and S. Elfar) and at the Laboratory of Sensorimotor Research for suggestions on previous versions of the manuscript, U. Schwarz for support during the collection of data for these experiments, and the National Institutes of Health Laboratory of Diagnostic Radiology Research for providing the MRIs of the monkeys.

Funding was provided by the National Eye Institute (R. H. Wurtz and G. H. Recanzone) and the National Research Council, National Institute on Deafness and Other Communication Disorders Grant DC-02371, The Klingenstein Fund, and the Sloan Foundation (G. H. Recanzone).

Address for reprint requests: R. H. Wurtz, Laboratory of Sensorimotor Research, Bldg. 49, Rm. 2A50, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892-4435.

Received 21 June 1999; accepted in final form 8 October 1999.

## REFERENCES

- ALLMAN, J., MIEZIN, F., AND MCGUINNESS, E. Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* 14: 105–126, 1985.
- BUSHNELL, M. C., GOLDBERG, M. E., AND ROBINSON, D. L. Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. *J. Neurophysiol.* 46: 755–772, 1981.
- DESIMONE, R. AND DUNCAN, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18: 193–222, 1995.
- DESIMONE, R. AND UNGERLEIDER, L. G. Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J. Comp. Neurol.* 248: 164–189, 1986.
- DURSTELER, M. R., WURTZ, R. H., AND NEWSOME, W. T. Directional pursuit deficits following lesions of the foveal representation within the superior temporal sulcus of the macaque monkey. *J. Neurophysiol.* 57: 1262–1287, 1987.
- FERRERA, V. P. AND LISBERGER, S. G. Attention and target selection for smooth pursuit eye movements. *J. Neurosci.* 15: 7472–7482, 1995.
- FERRERA, V. P. AND LISBERGER, S. G. Neuronal responses in visual areas MT and MST during smooth pursuit target selection. *J. Neurophysiol.* 78: 1433–1446, 1997.
- FUCHS, A. H. AND ROBINSON, D. A. A method for measuring horizontal and vertical eye movement chronically in the monkey. *J. Appl. Physiol.* 21: 1068–1070, 1966.
- GOLDBERG, M. E. AND WURTZ, R. H. Activity of superior colliculus in behaving monkey. II. Effect of attention on neuronal responses. *J. Neurophysiol.* 35: 560–574, 1972.
- GROH, J. M., BORN, R. T., AND NEWSOME, W. T. How is a sensory map read out? Effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements. *J. Neurosci.* 17: 4312–4330, 1997.
- HAENNY, P. E. AND SCHILLER, P. H. State dependent activity in monkey visual cortex. I. Single cell activity in V1 and V4 on visual tasks. *Exp. Brain Res.* 69: 225–244, 1988.

- JUDGE, S. J., RICHMOND, B. J., AND CHU, F. C. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res.* 20: 535–538, 1980.
- KOMATSU, H. AND WURTZ, R. H. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J. Neurophysiol.* 60: 580–603, 1988.
- KOMATSU, H. AND WURTZ, R. H. Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. *J. Neurophysiol.* 62: 31–47, 1989.
- KUSTOV, A. A. AND ROBINSON, D. L. Shared neural control of attentional shifts and eye movements. *Nature* 384: 74–77, 1996.
- LUCK, S. J., CHELAZZI, L., HILLYARD, S. A., 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.
- LYNCH, J. C., MOUNTCASTLE, V. B., TALBOT, W. H., AND YIN, T. C. Parietal lobe mechanisms for directed visual attention. *J. Neurophysiol.* 40: 362–389, 1977.
- MAUNSELL, J.H.R. AND VAN ESSEN, D. C. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation. *J. Neurophysiol.* 49: 1127–1147, 1983.
- MCADAMS, C. J. AND MAUNSELL, J.H.R. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* 19: 431–441, 1999.
- MORAN, J. AND DESIMONE, R. Selective attention gates visual processing in the extrastriate cortex. *Science* 229: 782–784, 1985.
- MOTTER, B. C. 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.
- MOTTER, B. C. Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J. Neurosci.* 14: 2178–2189, 1994a.
- MOTTER, B. C. Neural correlates of feature selective memory and pop-out in extrastriate area V4. *J. Neurosci.* 14: 2190–2199, 1994b.
- NEWSOME, W. T., WURTZ, R. H., DURSTELER, M. R., AND MIKAMI, A. Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *J. Neurosci.* 5: 825–840, 1985.
- NEWSOME, W. T., WURTZ, R. H., AND KOMATSU, H. Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *J. Neurophysiol.* 60: 604–620, 1988.
- RECANZONE, G. H., SCHWARZ, U., AND WURTZ, R. H. Responses of MT and MST neurons to one and two moving objects in the receptive field. *J. Neurophysiol.* 78: 2904–2915, 1997.
- RECANZONE, G. H., WURTZ, R. H., AND SCHWARZ, U. Attentional modulation of neuronal responses in MT and MST of a macaque monkey performing a visual discrimination task. *Soc. Neurosci. Abstr.* 19: 973, 1993.
- RECANZONE, G. H. AND WURTZ, R. H. Shift in smooth pursuit initiation and MT and MST neuronal activity under different stimulus conditions. *J. Neurophysiol.* 82: 1710–1727, 1999.
- REYNOLDS, J. H., CHELAZZI, L., AND DESIMONE, R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* 19: 1736–1753, 1999.
- RICHMOND, B. J. AND SATO, T. Enhancement of inferior temporal neurons during visual discrimination. *J. Neurophysiol.* 58: 1292–1306, 1987.
- ROBINSON, D. L., BOWMAN, E. M., AND KERTZMAN, C. Covert orienting of attention in macaques. II. Contributions of parietal cortex. *J. Neurophysiol.* 74: 698–712, 1995.
- SCHILLER, P. H. AND LEE, K. The effects of lateral geniculate nucleus, area V4, and middle temporal (MT) lesions on visually guided eye movements. *Vis. Neurosci.* 11: 229–241, 1994.
- SEIDEMANN, E. AND NEWSOME, W. T. Effect of spatial attention on the responses of area MT neurons. *J. Neurophysiol.* 81: 1783–1794, 1999.
- STEINMETZ, M. A., CONNOR, C. E., CONSTANTINIDIS, C., AND MCLAUGHLIN, J. R. Covert attention suppresses neuronal responses in area 7a of the posterior parietal cortex. *J. Neurophysiol.* 72: 1020–1023, 1994.
- TANAKA, K., FUKADA, Y., AND SAITO, H. A. Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol.* 62: 642–656, 1989.
- TANAKA, K., HIKOSAKA, K., SAITO, H., YUKIE, M., FUKADA, Y., AND IWAI, E. Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J. Neurophysiol.* 248: 190–222, 1986.
- TANAKA, K. AND SAITO, H. Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol.* 62: 626–641, 1989.
- TANAKA, K., SUGITA, Y., MORIYA, M., AND SAITO, H. Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. *J. Neurophysiol.* 69: 128–142, 1993.
- THIER, P. AND ERICKSON, R. G. Responses of visual-tracking neurons from cortical area MST-l to visual, eye and head motion. *Eur. J. Neurosci.* 4: 539–553, 1992.
- TREUE, S. AND MARTINEZ TRUJILLO, J. C. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399: 575–579, 1999.
- TREUE, S. AND MAUNSELL, J.H.R. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382: 539–541, 1996.
- TREUE, S. AND MAUNSELL, J.H.R. Effects of attention on the processing of motion in macaque visual cortical areas MT and MST. *J. Neurosci.* 19: 7591–7602, 1999.
- VAN ESSEN, D. C., MAUNSELL, J.H.R., AND BIXBY, J. L. The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *J. Comp. Neurol.* 199: 293–326, 1981.
- YAMASAKI, D. S. AND WURTZ, R. H. Recovery of function after lesions in the superior temporal sulcus in the monkey. *J. Neurophysiol.* 66: 651–673, 1990.