Neuronal Responses in Visual Areas MT and MST During Smooth Pursuit Target Selection

FERRERA, VINCENT P. and LISBERGER, STEPHEN G. Department of Physiology, W. M. Keck Foundation Center for Integrative Neuroscience, and Neuroscience Graduate Program, University of California at San Francisco, California, 94143

INTRODUCTION

A fundamental issue for both visual attention and voluntary movement control is the problem of selection. How do the systems that initiate voluntary movements or covert shifts of attention choose a particular target from an array of candidates? We have studied this problem with the use of voluntary smooth pursuit eye movements in monkeys as a model system. The smooth pursuit system in primates is a sensorimotor pathway that serves to stabilize the retinal image of moving targets. The pursuit system shows a high degree of selectivity in the way it transforms sensory input into motor output. One aspect of selectivity in the pursuit system is the ability to select one of several moving stimuli and to faithfully track the motion of the selected target while filtering out the motion of distractors (Ferrera and Lisberger 1995, 1997). The goal of this study is to understand how cortical motion processing in the middle temporal (MT) and middle superior temporal (MST) visual areas contributes to this selectivity.

The pursuit system uses signals derived from the retinal image motion of the desired target (Lisberger and Westbrook 1985; Lisberger et al. 1987). There is a great deal of evidence suggesting that these signals are provided by the cortical "motion pathway," which includes MT and MST. These areas are part of a pathway through parietal visual cortex that is primarily concerned with spatial and motion information (Albright 1984; Maunsell and Van Essen 1983a, b; Zeki 1978) and is thought to be involved in visual orienting and guidance (Andersen 1987; Goodale and Milner 1992; Ungerleider and Mishkin 1982). One of the major subcortical targets of MT and MST is the dorsolateral pontine nucleus (Boussaoud et al. 1992; Brodal 1978; Fries 1981; Glickstein et al. 1980, 1985; Maunsell and Van Essen 1983a; Ungerleider et al. 1984), which projects to the cerebellum and is thought to play a role in the generation of smooth eye movements (Brodal 1979, 1982; Gerrits and Voogd 1989; Mustari et al. 1988; Suzuki and Keller 1988; Suzuki et al. 1988; Their et al. 1988). Microstimulation of MT or MST during pursuit can affect smooth eye velocity (Groth et al. 1995; Komatsu and Wurtz 1989), whereas lesions of these areas lead to characteristic deficits in pursuit (Dursteler and Wurtz 1988; Dursteler et al. 1987).

Single-unit studies that have looked explicitly for possible roles of MT and MST in smooth pursuit have identified a population of "pursuit cells" that respond during smooth tracking of a small target in otherwise total darkness (Komatsu and Wurtz 1988a, b). These cells constituted ~30% of neurons in MT and MST and were distributed preferentially in regions where receptive fields (RFs) are near to or include the fovea (Desimone and Ungerleider 1986; Komatsu and Wurtz 1988a); pursuit cells were found in roughly equal proportions in foveal MT (MTf), ventrolateral MST (MSTl), and the dorsal pole of dorsomedial MST (MSTD). Some pursuit cells had a maintained response even when the target was stabilized on the retina or turned off briefly, thereby eliminating retinal image motion (Newsome et al. 1988). This "extraretinal" signal was strongest in MSTd, weak (but not absent) in MTf, and of variable strength in MSTl. At present it is not known to what extent the responses of pursuit cells represent signals related to eye position, eye velocity, or an abstract representation of the pursuit target in a nonretinocentric coordinate frame. Almost nothing is known about the role of these neurons in target selection. Do neuronal responses in MT and MST form an equally...
weighted representation of all moving stimuli or are they biased in a way that favors those stimuli that are to become the target of an eye movement? A more precise way to state this is to ask whether the response of MT or MST neurons is determined entirely by the retinal stimulus or whether it predicts the identity or movement of the intended target.

To address these issues, we trained two rhesus monkeys to perform a smooth pursuit target selection task. The monkeys were presented with two moving stimuli of different colors and were cued to track one stimulus or the other. We refer to the intended stimulus as the “target” and to the other as the “distractor.” We recorded from single neurons in MT and MST while the monkeys performed this task. The experiments were designed so that we could distinguish the purely visual (“retinal”) component of the response from the nonvisual (extraretinal) component. The extraretinal component contains information about which stimulus the monkey intends to look at. We sought to measure the magnitude and reliability of this extraretinal signal. We also sought to determine whether this extraretinal signal represents the direction or the color of the intended target. Finally, we were interested in the correspondence between neuronal response latency and the latency of the eye movement, because the latter has been found to vary systematically depending on the motion of the distractor (Ferrera and Lisberger 1995, 1997).

METHODS

Experiments were conducted on two male rhesus monkeys (Macaca mulatta) weighing ~6 and 7 kg. The same monkeys were used in earlier behavioral studies (Ferrera and Lisberger 1995, 1996). Our methods were approved by the UCSF Institutional Animal Care and Use Committee. Monkeys were trained to move voluntarily from their home cage to a primate chair. A method modified from Wurtz (1969) was used to train each monkey to attend a stationary task. Surgery was then performed under sterile conditions to implant a coil of wire on one eye (Judge et al. 1980) by 3 mm in a vertical position as the target. Target and distractor were separated by 10 mm in the vertical direction. The distractor axes of motion remained parallel.

Behavioral tasks

Monkeys were trained to track moving targets presented on a cathode ray tube monitor. We used a step-ramp target motion paradigm to minimize the occurrence of saccades during pursuit initiation (Rashbass 1961). Trials were initiated by requiring the monkey to look at a stationary central fixation light. After a short interval, a moving perifoveal target appeared. At the same time, the central fixation light was turned off and the monkey was required to track the target by initiating a smooth pursuit eye movement. The monkey was given a liquid reward provided gaze was kept directed toward the desired target for the duration of the trial. The monkey’s performance was monitored by tracking eye position relative to a ±3° fixation window centered around the target. Several steps were taken to ensure that the monkeys did not make anticipatory eye movements. First, the initial target location relative to the fixation mark was randomized from trial to trial. Second, the direction of target motion (toward or away from the vertical meridian) was also randomized. Third, trials were aborted if the monkey initiated an eye movement before the fixation light went off.

The task was one in which the monkey selected a pursuit target on the basis of a color cue (see Ferrera and Lisberger 1995) (Fig. 1). The monkey initially fixated on a small (0.4°) white square in the center of the screen. After a few hundred milliseconds, the white fixation mark was replaced by a colored square (red or green) of the same size and luminance. This was the cue. The color cue lasted for 250 ms and then changed back to the white (Fig. 1A). After a second time interval of 200 ms, the fixation target disappeared and two moving stimuli appeared (1 red, 1 green). This second interval was used so that neuronal responses to the onset of the pursuit targets would not be confused with any color-dependent response to the offset of the cue. The monkey’s task was to pursue the stimulus that matched the color of the cue (we call this stimulus...
the target and the nonmatching stimulus the distractor). After the stimuli appeared, the monkey initiated horizontal smooth pursuit with a latency generally in the range of 80–160 ms. On some randomly interleaved trials, only a single stimulus appeared so that we could measure the animal’s ‘normal’ pursuit under similar visual conditions. The color, initial position, and direction of the target were randomized trial-to-trial so that the monkey could not anticipate the direction of the required eye movement.

Figure 1B shows the display configuration. The dashed squares indicate potential target locations. These locations were offset both horizontally and vertically from the position of the initial fixation mark so that, to foveate the target, the monkey was required to make a vertical corrective saccade as well as horizontal pursuit. However, the latency of the saccade was usually much longer than that of pursuit. Target and distractor positions were randomized from trial to trial, with the constraint that the target and distractor never had the same vertical position. The distractor could move in the same direction as the target or in the opposite direction. To optimize neuronal responses, this basic configuration was modified by independently scaling and rotating the position and velocity of all display elements in polar coordinates. This allowed us to place at least one target inside a cell’s RF and to align the motion of the target with the cell’s preferred direction. The axes of motion of the target and distractor always remained parallel. When the axis of motion was close to horizontal, the initial target and distractor locations could be in the same or opposite hemifields. When the axis of motion was close to vertical, the target and distractor always started in opposite hemifields. For simplicity, we refer to the display as if it were in the standard configuration with all motion in the horizontal plane.

Trials were randomized within blocks so that the monkey was required to correctly complete exactly one trial of each type before proceeding to the next block of trials. On any given trial, the color of the target was randomized, as were its direction (toward or away from the vertical meridian) and position. The distractor condition (no distractor, same direction, or opposite direction) was also randomized. This resulted in a total of 48 trial types per block (2 colors × 2 directions × 4 positions × 3 distractor conditions). Trials in which the animal selected the wrong target (as judged by the direction of the corrective saccade) or initiated pursuit in the wrong direction were excluded from further analysis.

Visual stimulation

Pursuit targets were generated and controlled by a Univision Piranha video framebuffer with an on-board microprocessor (Texas Instruments TMS 34020). The output from the video board was displayed on a calibrated 20-in. color monitor (Barco) with a 60-Hz noninterlaced refresh rate. The monitor stood at a viewing distance of 30 in. so that the display area subtended roughly 30° horizontally by 20° vertically. The spatial resolution of the display was 1,280 pixels × 1,024 lines, and the depth was 8 bits per pixel. Pursuit targets were small (0.9°) colored squares presented on a uniform gray background. The luminance of the fixation mark, cue, and targets was 15.0 cd/m², whereas the background was 5.0 cd/m².

The framebuffer was programmed to send out digital pulses (frame sync) for timing purposes at the beginning of each frame in which a target was turned on or started to move. These pulses were recorded by the computer and stored along with the eye movement and spike data. We did not compensate for the delay between the occurrence of the frame sync and the actual drawing of the targets, which could be as long as 15 ms, but was typically ~8 ms. This delay affected neuronal and behavioral measurements equally.

Eye movement recording

Eye position was monitored with the use of a monocular scleral search coil system (CNC Engineering). Separate horizontal and vertical eye position signals were fed through an analog differentiator (low-pass, −3 dB at 25 Hz) to yield horizontal and vertical eye velocity. The eye position and eye velocity signals were then digitally sampled by computer at 1 kHz per channel and stored on disk for further analysis.

Neuronal recording and data collection

When behavioral training was complete, a recording chamber (20 mm diam) was implanted on the skull overlying the superior temporal sulcus (STS) or the operculum. We used two approaches to MT and MST: a horizontal approach through V1 with the recording chamber angled 15° from the horizontal plane (monkey B) and a vertical approach through the inferior parietal lobe (monkey F). In both cases, the center of the recording chamber was positioned 15 mm lateral to the midline. Each day a hydraulic microdrive was mounted on the recording chamber, filled with sterile mineral oil, and sealed. Transdural recordings were made with the use of Pt/Ir or tungsten extracellular electrodes with impedances of ~1–2 MΩ at 1 kHz (Wolbarsht et al. 1960). Signals from the microelectrode were amplified, filtered, and monitored on an oscilloscope and audio monitor.

The animals performed one of several eye movement tasks while we searched for units. Units were isolated on the basis of waveform, with the requirement that the peak of the action potential be ≥2 times the peak of the background noise. Action potentials were converted to digital pulses with a time-amplitude window discriminator. The time of occurrence of action potentials was recorded with a precision of 0.01 ms on the same time base used for recording vertical sync pulses from the video board. When a unit was isolated, stimulus parameters such as the axis and speed of target motion were adjusted to optimize its response. Data were then collected while the animal performed the target selection task. When data collection ended for the target selection tasks, we attempted to map the RF of the unit by hand while the monkey fixated a stationary spot.

Verification of recording sites

For each monkey, we initially sampled over a wide area of the STS, effectively ‘mapping’ the RF topography of the posterior and anterior banks. These RFs were used to guide electrode penetrations and to help with the assignment of individual units to MT or MST. However, not every RF corresponds to a single unit used in the study. Many RF plots correspond to multiunit recording sites and some correspond to sites that were not used because their RFs were too eccentric or they did not respond to small moving targets. It should be noted that many sites were excluded because the RFs were too large to plot within the confines of our tangent screen, which was limited to the central 25° of the visual field (as measured along the horizontal meridian). Figure 2 shows the relationship between RF size (square root of RF area) and eccentricity for single units and multiunit sites that were assigned to MT (n = 76) or MST (n = 36). We calculated linear regression lines for RF size versus eccentricity, and the slopes of these lines are in general agreement with other work (Komatsu and Wurtz 1988a). We subsequently limited our recordings to regions of MT and MST representing the central 10°, which is where pursuit initiation is optimal. We found two distinct representations of the central visual field in the STS. One was located more dorsally on the anterior bank and corresponded to MSTd. The other was located more ventrally on the posterior bank and corresponded to MTf or MSTf. We were not able to distinguish reliably between MTf and MSTf. In the
RESULTS

Responses to single targets

Before assessing the role of MT and MST in target selection, it is useful to determine whether responses were appropriate for identifying the direction, color, and position of the moving stimuli. We did this by extracting responses to single targets that were presented on random trials during the course of the target selection task. We constructed peristimulus time histograms for all cells that provided sufficient data (28 in MT, 83 in MST). These histograms are depicted in Fig. 3. Data for MT and MST were combined. Figure 3A shows responses at the preferred location, whereas Fig. 3, B–D, shows responses at nonpreferred locations. The black histograms are the responses to the preferred direction or color and the gray histograms are responses to the nonpreferred direction or color. For each cell, the trial-by-trial responses were first summed and then normalized to the number of trials. Then the responses for all cells were summed and normalized by the number of cells. The Y-axis therefore represents the firing rate per cell per trial. The X-axis represents time in milliseconds. Histograms were smoothed with the use of an exponential low-pass filter with a time constant of 10.0 ms. The bar on the X-axis demarcates the first 100 ms of target motion. Smooth pursuit eye movements were generally initiated after the end of this 100-ms interval.

Figure 3 shows that the net response in MT and MST was selective for stimulus location and direction and that the response latency was shorter than the average latency of pursuit by about half. The latency of neuronal responses in this study was much shorter than that reported by Newsome et al. (1988), who recorded mainly from neurons that started to respond well after the initiation of pursuit. This could be due to different stimulus conditions in the two studies or to the possibility that we isolated different subpopulations of neurons. Direction selectivity is apparent at all stimulus locations. From the position sensitivity, one can infer that two discrete moving stimuli will activate topographically distinct but overlapping populations of neurons. Color selectivity was poor, although it should be noted that we made no attempt to measure cell’s true color preferences, which would entail the use of a much broader spectrum of target colors. When we compared responses to the green and red stimuli on a cell-by-cell basis, we found no systematic preference for one color or the other (MT: green average 29.1 spikes/s, red average 21.6 spikes/s; paired t-test, green vs. red $P = 0.49$; MST: green average 30.5 spikes/s, red average 29.1 spikes/s, paired t-test, $P = 0.13$). This indicates that the small color preference shown in Fig. 3 probably was not contaminated by a luminance artifact that systematically biased responses for one color or the other.

Although Fig. 3 does not contain any fundamentally new information about neuronal responses in the STS, it does...
help to verify that we have recorded from the relevant population of cells within MT and MST, i.e., those that have a direction-selective response before the initiation of pursuit. These are the cells that are most likely to show cognitive signals related to target selection.

Selection bias in MT and MST

We recorded from MT and MST neurons with the use of a task designed to dissociate the sensory and cognitive components of the neuronal response. Figure 4 shows the responses of a single neuron recorded in MST to a subset of trials for this task. Each column shows responses to the visual stimulus configuration depicted at the top of the column. Filled squares represent the green stimulus and open squares represent the red stimulus. The cell had a large RF that covered both stimuli. Each row corresponds to a particular instruction about which stimulus to track (green or red). Each of the eight subplots shows the spike histogram (unit), stimulus onset (stim), and horizontal eye velocity (E\textsuperscript{+}), respectively. The eye velocity traces have been shifted vertically to fit in the space provided. The beginning of each trace represents zero eye velocity. Positive (upward) deflections correspond to rightward movements and negative (downward) deflections correspond to leftward movements. This cell had both an early and a late response. The vertical dashed lines demarcate the time of the early response. The late response was highly selective for leftward image motion and was strongest during rightward eye movements. The early response (indicated by the vertical dashed lines) was triggered by the retinal stimulus but was significantly stronger when the animal was instructed to track the green target (unpaired t-test, \( P < 0.05 \)). The early response averaged 46.3 spikes/s on “track green” trials, whereas the response on “track red” trials averaged 29.5 spikes/s. This cell was typical of those that showed significant extraretinal response modulation.

The cell in Fig. 4 demonstrates another feature that is common among MT and MST neurons in that the cell responded both to the initial stimulus motion before pursuit and to the image motion that occurred during pursuit maintenance. This suggests that the cell responded to any motion inside its RF and did not distinguish between target and background motion. We also found many cells that responded only to the initial target motion and others that responded only to image motion during pursuit. We did not attempt to classify cells on this basis because we felt that the responses before and during pursuit could probably be predicted from the cell’s velocity tuning, RF location, and spatial summation within the RF (Born and Tootell 1992).

Because of the need to control the retinal stimulus, and also because we are interested in the events that precede pursuit initiation, we restricted ourselves to looking at neural activity that occurred after the appearance of the stimuli and before the initiation of the eye movement. For each neuron, we computed the average firing rate during the interval from 30 to 100 ms after the onset of stimulus motion (Fig. 4, ---). The shortest neuronal latencies in MT are \( \sim 40 \) ms (Maunsell 1987), so that eye movements with latencies as short as 60 ms should not affect neuronal responses during the first 100 ms of stimulus motion. Each cell was tested with all combinations of stimulus color, direction, and position. From these we selected the stimulus configuration that gave the best overall response. We then took all the responses to the best stimulus configuration and sorted them according to the color of the cue that instructed the monkey as to which of the two stimuli to track. Figure 5 shows, for each cell, the responses to the best stimulus configuration sorted as a function of the instruction. We ran each cell through two tests. The first test checked that the selection bias for the best stimulus was consistent with the selection bias averaged over all stimuli. We then performed a t-test on the responses to the best stimulus (unpaired, 1-tailed, \( P \leq 0.05 \)). Cells that passed both tests are plotted as filled circles. Cells that passed only the consistency test are plotted as open circles and cells that passed neither test are plotted...
as small squares. Although the selection bias is generally rather weak, both areas have a substantial number of cells that passed both tests (14% in MT and 26% in MST).

Figure 5 shows responses to the best overall stimulus configuration, regardless of whether the distractor and target moved in the same or opposite directions. However, there is reason to believe that the nature of the task may differ in the two conditions; when target and distractor are moving in opposite directions, the monkey must choose both the color and direction of the target, and when target and distractor are moving in the same direction, the monkey only needs to choose the color. We therefore repeated our analysis for the two distractor conditions separately. For each distractor condition we determined the stimulus that gave the best overall response and then we tested whether there was a significant difference between the average response to that stimulus and the gray histogram is the response following the preferred instruction. For each cell, the trial-by-trial responses were first summed and then normalized to the number

\[ \text{MI} = \frac{|R_g - R_r|}{(R_g + R_r)} \]

where \( R_g \) is the response following a green cue and \( R_r \) is the response following a red cue. We computed an index for the overall response and for each distractor condition (same/opposite) separately. The average modulation index was 0.2 (MT: 0.2 ± 0.02, median = 0.14; MST: 0.2 ± 0.02, median = 0.14). The average modulation index for units with significant effects was in the range of 0.25–0.35 (MT: 0.34 ± 0.07, median = 0.35; MST: 0.32 ± 0.05, median = 0.26). Results for the same and opposite distractor conditions were similar. The cell in Fig. 4 had a modulation index of 0.22.

To get a sense of the time course of the selection bias, we constructed cumulative peristimulus time histograms for the 30 cells that showed significant effects in Fig. 5. These histograms are plotted in Fig. 6. Data for MT and MST were combined. On the left are the plotted cells that responded more strongly when the instruction was to track the green target, and on the right are the cells that responded more strongly when the instruction was to track red. The black histograms are the responses following the preferred instruction and the gray histograms are the responses following the nonpreferred instruction. For each cell, the trial-by-trial responses were first summed and then normalized to the number

\[ X \text{-axis: } 1 \text{st } 100 \text{ ms of target motion. Histograms are plotted in Fig. 6.} \]

The selection bias index was calculated as the proportion of cells with significant differences. A: selection bias in MT. Fraction in the bottom right corner: proportion of cells with significant differences. B: selection bias in MST.
of trials. Then the responses for all cells were summed and normalized by the number of cells. The Y-axis therefore represents the firing rate per cell per trial. The X-axis represents time with a resolution of 1 ms. The bar on the X-axis demarcates the first 100 ms of target motion. The selection bias is the difference between the black and gray histograms and is plotted at bottom. The peak selection bias is about one-third of the peak response and is present from the very beginning of the response. Because we selected the most significant cells and sorted them in such a way as to show the maximal effect, Fig. 6 should be viewed as an upper bound on the selection bias that is present in MT and MST.

The data in Figs. 4–6 raise the issue of what attribute is selected when the animal selects a target. In these experiments, the selected attribute could be either the color or the motion of the target. Because we tested cells with all combinations of 2 colors × 2 directions, we can look for effects of both color and direction to determine which attribute of the intended target is predicted by the response of a given cell. For the MST cell in Fig. 4, it is clear that the early response is consistently stronger when the animal is cued to track green (compare top and bottom rows), but that there is little correlation with the direction of the target.

To determine whether the extraretinal response was better correlated with the color or direction of the intended target, we sorted responses first according to the color of the target (averaging over all positions and directions of movement) and then according to target direction (averaging over position and color). We used the responses for all trials with two stimuli moving in opposite directions. This means that instead of looking at responses to the best stimulus configuration, as in Fig. 5, we are looking at responses to a group of stimuli, including many that are nonoptimal. This analysis is only valid for conditions in which the target and distractor move in opposite directions, so that the direction of pursuit is not completely determined by the stimulus. It is also important to have the same number of trials for each condition, so that when we sort responses by target color and direction, each of the sorted categories has the same net stimulus motion and color. The red-green and right-left differences should be zero for any cell that has a purely sensory response. Nonzero differences indicate the extent to which the cell’s response is predictive of the color or direction of the intended target. In Fig. 7 the red-green difference is plotted against the left-right difference for both MT and MST. (“Left-right” refers to the stimulus in the standard configuration, as described in METHODS. The actual axis of motion was rotated to match the preferred direction of each cell.) The overall distribution is somewhat elongated along the right-left axis, indicating that the selection bias in MT and MST is more predictive of target direction than target color. A two-way analysis of variance showed that of 28 cells in MT, 3 had a significant ($P \leq 0.05$) effect for color and 1 had a significant effect for direction. For 83 cells in MST, 4 were significant for color and 6 were significant for direction. The statistical significance of the effects is diluted by the fact that responses were averaged over optimal and nonoptimal stimuli, thus increasing the response variance.

**Relationship between neuronal and behavioral latency**

We have previously shown that the presence of a moving distractor shifts the latency of pursuit (Ferrera and Lisberger 1995, 1997). Distractors moving in the same direction as the target reduce the latency of pursuit, whereas distractors moving in the opposite direction increase latency. The latency shift can be interpreted as evidence of a competitive mechanism for target selection. We were therefore motivated to examine the relationship between neuronal response latencies and the latency of smooth pursuit. We did this on a cell-by-cell and trial-type-by-trial-type basis by taking all trials of a given type and constructing average spike histograms and eye velocity traces. Only trials in which the target moved toward the fovea were included. The neuronal and behavioral latencies were determined by visual inspection of the averaged responses: the action potential peristimulus time histograms for neuronal latency and the horizontal and vertical eye velocity traces for behavioral latency. Most cells had fairly crisp and reliable responses, so that there was no difficulty in determining the onset of the response. Cells that did not have clear onsets for targets moving toward the fovea were excluded from this analysis. The analysis was performed on 38 MT cells and 41 MST cells. For each individual cell that was used, neuronal latencies were taken only from trial types that had clear response onsets; thus responses to targets moving in the null direction were often excluded. Behavioral latencies were recorded for every trial type. In addition, neuronal responses were excluded if they occurred later than 200 ms or were clearly related to retinal image motion subsequent to the initiation of pursuit.

Scatterplots and distributions for neuronal and behavioral latencies are shown in Fig. 8. The scatterplots show only behavioral latencies that had an accompanying neuronal latency, but the distributions show all behavioral latencies. For a given neuron there was very little variation in latency across trial types. The individual data points for each cell occupy a narrow vertical column. Much of the variance in overall neuronal latency is therefore due to between-cell differences. The average neuronal latency in MST was
neuronal and behavioral latencies with the number of observations in parentheses. Table 1 also gives the population correlation coefficients (r) for neuronal versus behavioral latency, sorted by distractor condition, and the mean ± SE of the distribution of cell-by-cell correlations and regression slopes. The main finding is that there was a negligible correlation between neuronal latency and behavioral latency on both a cell-by-cell and a population basis. In fact, one could say these data show a clear dissociation between the responses of neurons in MT and MST and the behavioral output of the pursuit system.

**Relationship between neuronal response amplitude and behavioral latency**

It is possible that behavioral latency might be correlated not with neuronal latency but neuronal response amplitude. This would be the case if there were a firing rate threshold that needed to be reached before an eye movement was initiated (Hanes and Schall 1996). If counterdirectional movement suppresses neuronal responses in MT and MST, as suggested by others (Qian and Andersen 1994; Snowden et al. 1991), this could account for the longer behavioral latencies in the distractor-opposite condition. We tested this by comparing neuronal responses to different distractor conditions (single, same, and opposite). For each cell, we first found the single-target condition (direction and position) that gave the strongest response. We then computed the response to that target when it was paired with a distractor moving in the same direction and when paired with a distractor moving in the opposite direction. We performed this analysis for the same set of cells shown in Fig. 4. In Fig. 9 the response to paired targets (○, same condition; ●, opposite condition) showed very little dependence on trial type; however, behavioral latencies showed the characteristic dependence on distractor condition that we have documented previously. Specifically, the shortest behavioral latencies were found for trials in which the distractor moved in the same direction as the target, and the longest latencies were found for trials in which the distractor moved in the opposite direction. Single-target latencies fell somewhere in between.

We calculated correlation coefficients for the entire population of neuronal and behavioral latencies. We also calculated correlation coefficients and regression lines on a cell-by-cell basis. For regression, we used neuronal latency rather than behavioral latency as the dependent variable to avoid regression lines with infinite slope. Table 1 shows average neuronal latencies, sorted by distractor condition (●, target and distractor same direction; ○, target and distractor opposite directions). Scatterplot shows only data for trial types that provided both a neuronal and a behavioral latency measurement. Distributions below scatterplot: neuronal latencies sorted by distractor condition, with latencies for single targets in gray. Distributions at right of scatterplot: behavioral latencies for all trial types, again sorted by distractor condition with single targets in gray. B: data for MST plotted in the same format. Horizontal and vertical scales for both the scatterplots and the distributions are identical. All latencies were taken from averaged histograms with 4-ms time bins. For display purposes only, the data points in the scatterplots were randomly jittered by ±2 ms in both dimensions to reduce overlap among points falling in the same bin.

![Neuronal Latency](image)

**Fig. 8.** Neuronal and behavioral response latencies. A: MT neuronal vs. behavioral latency sorted by distractor condition (+, single targets; ○, target and distractor same direction; ●, target and distractor opposite directions). Scatterplot shows only data for trial types that provided both a neuronal and a behavioral latency measurement. Distributions below scatterplot: neuronal latencies sorted by distractor condition, with latencies for single targets in gray. Distributions at right of scatterplot: behavioral latencies for all trial types, again sorted by distractor condition with single targets in gray. B: data for MST plotted in the same format. Horizontal and vertical scales for both the scatterplots and the distributions are identical. All latencies were taken from averaged histograms with 4-ms time bins. For display purposes only, the data points in the scatterplots were randomly jittered by ±2 ms in both dimensions to reduce overlap among points falling in the same bin.

<table>
<thead>
<tr>
<th>Table 1. Relationship between neuronal and behavioral latency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MT</strong></td>
</tr>
<tr>
<td>Neuronal latency</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>Single</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Opposite</td>
</tr>
<tr>
<td>Behavioral latency</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>Single</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Opposite</td>
</tr>
<tr>
<td>Neuronal vs. behavioral latency cell by cell</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Correlation (r)</td>
</tr>
<tr>
<td>Regression slope</td>
</tr>
<tr>
<td>Population correlations (r)</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>Single</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Opposite</td>
</tr>
</tbody>
</table>

Values for neuronal vs. behavioral latency are means ± SE. Numbers in parentheses: number of observations. n, number of cells. MT, middle temporal visual area; MST, medial superior temporal visual area.
predict that the presence of a distractor should shorten the latency of pursuit (compared with a single target) regardless of the motion of the distractor relative to the target. This prediction also contradicts the observed behavioral pattern. Overall, it appears that neither the latency nor the amplitude of responses in MT or MST can account for the effect of a moving distractor on pursuit latency.

Estimation of the ‘net’ directional signal

As a final measure, we were interested in determining the net directional signal that is available to stages of the pursuit pathway downstream from MT and MST. In other words, what does the activity in MT and MST look like to areas that are “reading out” the distributed representation of image motion? In Fig. 10 we show the summed activity of all MT and MST neurons (the histograms were constructed as in Figs. 3 and 6) when the target moved in either the preferred direction or a null direction.

We performed a variant of this analysis in which, rather than using the target that gave the best response, we used the responses to all stimuli. In other words, we sorted the responses by distractor condition (single, same, or opposite), but within each condition we averaged over color, direction, and position. The results were different in two ways. First, responses on the whole were weaker because they included many nonoptimal stimuli. Second, responses to paired targets were generally stronger than responses to single targets. The mean firing rates in spikes/s were as follows: MT single = 21.5, same = 30.0, opposite = 29.4; MST single = 30.0, same = 36.5, opposite = 36.5. This is predictable given that when two targets were present there was a greater chance that at least one of them would be close to optimal. Thus the total amount of activity in MT and MST is greater for two targets than for a single target. This might lead one to
(black) or null (gray) direction. In Fig. 10, A and B, we included responses to stimuli comprising a target and distractor where the target was presented at the preferred location for each cell. In this case, there is a robust directional signal regardless of whether the target and distractor move in the same (Fig. 10A) or opposite (Fig. 10B) direction. This confirms the results of Fig. 9, which also show that there is little net influence of distractors at nonpreferred locations.

Next, we looked at responses to stimulus configurations in which either the target or distractor could be at the preferred location (Fig. 10, C and D). The black histograms are responses when the target moved in the preferred direction and the gray histograms are responses when the target moved in the null direction. If, on any given trial, the target and distractor activate distinct populations of neurons (as Fig. 3 implies), then Fig. 10, C and D, represents the summed activity of those two populations. In Fig. 10C it makes no difference whether the target or distractor is at the preferred location because they both move in the same direction. However, in Fig. 10D, the target and distractor moved in opposite directions and the activity they created nearly cancels. What is left is a small residual response indicating the direction of the target.

**DISCUSSION**

The main goal of this study was to determine how multiple discrete moving targets are represented in MT and MST and whether that representation is biased in a way that predicts which target will be used to guide smooth pursuit. For MT, it is known that there is a crude topographic map of retinal position (Maunsell and Van Essen 1983a,b) combined with an orderly representation of direction (Albright 1984). One might therefore imagine that two moving targets should be represented by two peaks of activity in MT corresponding todirection columns at different topographic locations. Ideally, one would like to compare the amplitude of these two peaks to find out whether the peak corresponding to the distractor is smaller than that corresponding to the target. However, because we could only measure the activity at a single site, we instead measured the response to the same retinal stimulus while varying the instruction (cue) given to the animal regarding which target to track. We were thus able to dissociate the response of a single neuron into one component that was attributable to sensory (retinal) signals and another that was attributable to cognitive (extraretinal) factors. We refer to the latter as the selection bias or predictive component of the response.

We found that the responses of some cells (14% in MT, 26% in MST) reliably predicted the identity of the intended target. The strength of the predictive component for units with significant effects represented roughly a 25–35% modulation of response. The full task comprised trials with all permutations of target color, direction, and position. To measure selection bias, we compared responses to the stimulus configuration that yielded the optimal response sorted by the color of the cue. However, for a single-stimulus configuration, it is ambiguous whether the selection bias predicts the motion or the color of the intended target. To sort this out, we looked at responses to all stimuli in which the target and distractor moved in different directions. We sorted these responses both according to the color and the direction of the target. The pattern that emerges is shown in Fig. 7. It suggests that cells in MT and MST are somewhat better predictors of target direction than target color. (However, it should be noted that although every attempt was made to find the preferred axis of motion for each cell, no attempt was made to find the preferred color.)

An inevitable question is why we used a color-based selection task to examine cortical areas that have little conventional color selectivity. We used the color task because we found that it was the easiest way to train the monkeys. After training the monkeys to perform the color task, we attempted to train them to use a direction cue. The direction cue was a small patch of moving random dots presented just before the appearance of the target and distractor. The direction of the dots specified the direction of the desired target. Both monkeys were able to achieve only slightly better than chance performance on the direction-cue task. We concluded that target selection based on color is a more natural task than target selection based on motion. At some level, the intention to look at a target of a particular color must be linked with the direction of the target to specify the correct eye movement (Treisman and Gelade 1980). Therefore it is reasonable to expect that selection effects should be found in areas that represent either the color or the motion of the target (or both). Some of the selection bias found in MT and MST seemed to be related to color of the target rather than its direction (e.g., Fig. 4). The color-dependent component of the selection bias might reflect attentional modulation in cortical areas that are more sensitive to color and that have connections with MT and MST, such as V4. Our single-unit results with the use of the color task are in general agreement with another study in which monkeys were cued for the shape or location of the target (Recanzone et al. 1993). Our results are also consistent with a report that color can enhance neuronal sensitivity to motion in MT (Crone and Albright 1996).

The response modulation we observed in MT and MST is consistent with a “biased competition” model for target selection (Desimone and Duncan 1995). The notion of attention as an outcome of competition among sensory signals goes back at least as far as Broadbent (1958) and has been employed in various computational models of attention and eye movements (Ferrera and Lisberger 1995, 1996; Koch and Ullman 1985; Scheinberg and Zelinsky 1993). One version of such a model is shown in Fig. 11. In this model, target and distractor velocity signals are processed by a competitive network and the output of this network is the desired eye velocity command used by the motor system to generate the eye movement. A selection bias that favors the direction of the target determines the output of the network. The current results indicate that the predictive component of responses in MT and MST might be a correlate of the selection bias in the model.

What accounts for the behavioral selectivity of pursuit?

The strength of the selection bias we measured raises some interesting practical issues in relating neuronal selectivity to behavioral selectivity. One can imagine two extremes of
selectivity, which we will refer to as “strong selection” and “weak selection.” These correspond respectively to the “winner-take-all” and “vector-averaging” dichotomy suggested by others (Groh et al. 1995; Salzman and Newsome 1994). Under conditions similar to those of the present experiments, pursuit behavior shows strong selection. The direction of pursuit is identical to the direction of the selected target and is not influenced by the motion of the distractor (Ferrera and Lisberger 1997). It should be noted that when there is no attentional bias, the direction of pursuit is intermediate between that of the target and distractor (see Groh et al. 1995; Lisberger and Ferrera 1996). However, the effects of attention that have been documented in visual cortex, including those of the present study, clearly indicate weak selection. These effects are almost always limited to less than half of the cells examined, typically ~25–40%. In the cells that show significant attentional modulation, the mean effect is usually something like 30% enhancement or suppression of the response. This leaves a considerable gap between cortical physiology and eye movement behavior. To bridge this gap, we have proposed that there is a competitive network that serves to amplify the weak selection bias, converting it into an all-or-none output. It has been shown mathematically that some types of competitive networks can generate winner-take-all output from an arbitrarily small selection bias (Yuille and Grzywacz 1989), so that weak cortical selection is not necessarily an obstacle for the generation of accurate movements.

It is worth considering whether the task we used could be manipulated to obtain stronger effects. Other studies have suggested that attentional effects are strongest when both target and distractor fall within the RF of the cell in question (Connor et al. 1996; Moran and Desimone 1985; Treue and Maunsell 1996). Although we did not explore this systematically, we kept the separation between the target and distractor in the range of 3–12°. Figure 3 shows that there are substantial responses to stimuli presented at nonpreferred locations. From this it may be inferred that most stimulus locations were inside the RF of a given cell. In addition, behavioral measurements indicate that distractor motion affects the latency of pursuit as long as the distractor is within 16° of the target (Ferrera and Lisberger 1996) and that it makes no difference if the target and distractor are in the same or opposite hemifields (Ferrera and Lisberger 1995). These considerations suggest that the neural site of target selection should have RFs or horizontal connections that are bilateral and cover at least the central 16° of the visual field, like those in MSTd or dorsolateral pontine nucleus (Suzuki et al. 1990). We have also observed that pursuit performance tends to break down at small separations. When the target and distractor are separated by <3°, animals make substantially more mistakes and latencies are exceptionally long, suggesting that the pursuit system has difficulty resolving motion signals within small RFs, such as those in MT.

Another factor that works against obtaining strong effects is that we only looked at responses during the first 100 ms of target motion. Full attentional effects might take longer to develop (Duncan et al. 1994; Preddie et al. 1995). We considered manipulating the task in a way that would extend pursuit latencies beyond their normal range, but we felt that this would have distorted our results. Effects of attention that occur on longer time scales might not be relevant to pursuit initiation under natural conditions, where response latencies are on the order of 100 ms.

What accounts for the behavioral latency of pursuit?

Behavioral latencies for smooth pursuit show a characteristic dependence on distractor motion that can be seen in Fig. 8 and Table 1. However, we found no systematic relationship between behavioral latency and either neuronal response latency or amplitude. The dissociation between responses in MT and MST and the latency of pursuit is interesting because others have found that manipulations of the visual stimulus can lead to correlated changes in neuronal and behavioral latency. For example, Krauzlis (1991) found that varying the “motion onset delay” of a visual target (i.e., the amount of time the target remains stationary before it begins to move in the step-ramp paradigm) affects pursuit latency such that increasing the delay leads to shorter latencies. Movshon et al. (1990) subsequently showed that increasing motion onset delay also reduces neuronal latencies in MT in anesthetized monkeys. Kawano et al. (1994) showed a similar correlation in alert monkeys. They found that increasing the speed of a target reduces the latency of ocular following (a smooth eye movement evoked by movements of a large-field stimulus) and also reduces the latency of neuronal responses in MST. In fact, the slope of the function relating neuronal to behavioral latency was nearly 1.0, indicating a very close temporal correspondence between responses in MST and the generation of ocular following. For visuomotor behaviors such as smooth pursuit, visual processing delays necessarily result in delayed motor responses. By showing a dissociation between visual responses in cortex and behavioral latency shifts induced by competing stimuli, we can argue that these behavioral effects are not due to sensory processing delays.

These observations join a growing body of evidence suggesting that activity in MT and MST cannot quantitatively account for smooth pursuit behavior without explicit assumptions about downstream mechanisms (Kiorpes et al. 1996). One candidate for such a mechanism is the competitive network mentioned earlier. In previous work, we have shown that distractor-induced shifts in pursuit latency are consistent with the internal dynamics of a competitive net-
work (Ferrera and Lisberger 1995, 1997). The lack of a correlation between neuronal response latency or amplitude and behavioral latency indicates that neither MT nor MST is the site of this competitive bottleneck.

Another class of models that might account for the behavioral data is that of statistical decision theory. These models work by “weighing the evidence” in favor of one direction or the other and making a decision based on a statistical criterion (e.g., Carpenter and Williams 1996; Shadlen et al. 1996). The “evidence” in this case is the directional bias remaining after one averages out the confounding effects of distractor motion. If one assumes that the pursuit system uses a weighted sum of activity in MT and MST (each neuron’s response weighted by its preferred direction), then the direction bias available to downstream stages of the pathway is roughly equivalent to the differential response to motion shown in Fig. 7 and Fig. 10, C and D. Figure 10C shows that when the target and distractor move in the same direction, there is a strong directional signal. But when target and distractor move in opposite directions (Fig. 7 and Fig. 10D), the net directional signal is many times weaker. The directional signal in Fig. 10C also has much greater statistical reliability than that in Fig. 10D. The greater magnitude and reliability of the direction signal should translate into faster responses when the target and distractor move in the same direction.

This explanation works if downstream stages of the pursuit pathway simply sum all the activity in MT and MST. However, the explanation fails if later stages are able to selectively access activity that is related to the target rather than activity related to the distractor. In this case, Fig. 10, A and B, shows that the directional information about the target is equally robust regardless of the relative motion of target and distractor, so there should be no effect on pursuit latency.

Theoretical considerations

The current results are surprising because a number of theoretical arguments and experimental observations gave reason to expect much stronger response modulation in MT and MST during target selection. Recently, Treue and Maunsell (1996) reported robust modulation of responses in MT and MST when monkeys performed a discrimination task requiring covert attention to moving targets. If there is a necessary link between covert shifts of attention and the programming of eye movements, as postulated by Rizzolati et al. (1987), then one would predict robust attentional effects during eye movement tasks in cortical areas that are modulated by covert attention. Another theory, known as the “what/where” or “action/perception” hypothesis, predicts that attentional effects in the parietal pathway, including MT and MST, should be stronger for visual guidance tasks than for covert attention tasks (Goda 1995; Ungerleider and Mishkin 1982). The observation of weak modulation during smooth pursuit target selection contradicts aspects of both theories and indicates that there might be a dissociation between the effects of covert attention and eye movements in MT and MST. This complements the observation that there is selective enhancement of responses in the frontal eye field preceding saccades (Goldberg and Bushnell 1981) but not during covert attentional shifts. Bushnell et al. (1980) also found both attention-related and saccade-related enhancement in area 7a of parietal cortex. The overall pattern of results favors the idea that there are distinct but overlapping cortical substrates for overt and covert orienting, with prefrontal cortex being more involved in overt shifts and parietal cortex more involved in covert shifts of attention. This is consistent with the view that there are distinct mechanisms for visual attention and motor response selection (Pashler 1991).

Other observations have revealed that pursuit cells in MT and MST (i.e., cells that have RFs near the fovea and that respond during the maintenance phase of pursuit) have visual responses that are supplemented by an extraretinal signal possibly related to eye position or velocity (Newsome et al. 1988). On the other hand, cells in MST with RFs away from the fovea have visual responses that are suppressed during maintained pursuit (Erickson and Thier 1991). It now seems likely that MST (and to a lesser extent, MT) is involved in distinguishing the “true” motion of the target from image motion of the stationary background that is induced by the eye movement, a distinction that is important for perceptual stability. However, even though MT and MST respond vigorously to stimulus motion before the initiation of pursuit, they appear not to have a prominent role in the decision that determines which of several moving stimuli will ultimately become the target of a smooth eye movement. Cognitive signals in MT and MST therefore appear to be more closely related to pursuit maintenance than to pursuit initiation.

We thank Dr. Karl Gegenfurtner for sharing graphics software and Dr. John Maunsell for helpful comments.

This work was supported by the McDonnell-Pew Program in Cognitive Neuroscience Fellowship JSMF 92-38 to V. P. Ferrera and National Eye Institute Grant EY-03878 to S. G. Lisberger.

Present address and address for reprint requests: V. P. Ferrera, Dept. of Psychiatry and Center for Neurobiology and Behavior, Columbia University, 722 West 168th St., New York, NY 10032.

Received 13 December 1996; accepted in final form 21 May 1997.

REFERENCES


Bushnell, M. C., Goldberg, M. E., and Robinson, D. L. Behavioral en-


