Extrastriate Area MST and Parietal Area VIP Similarly Represent Forward Headings

James B. Maciokas1 and Kenneth H. Britten1,2
1Center for Neuroscience and 2Department of Neurobiology, Physiology, and Behavior, University of California, Davis, California

Submitted 10 December 2009; accepted in final form 22 April 2010

Maciokas JB, Britten KH. Extrastriate area MST and parietal area VIP similarly represent forward headings. J Neurophysiol 104: 239–247, 2010. First published April 28, 2010; doi:10.1152/jn.01083.2009. Many studies have documented the involvement of medial superior temporal extrastriate area (MST) in the perception of heading based on optic flow information. Furthermore, both heading perception and the responses of MST neurons are relatively stable in the presence of eye movements that distort the retinal flow information on which perception is based. Area VIP in the posterior parietal cortex also contains a robust representation of optic flow cues for heading. However, the studies in the two areas were frequently conducted using different stimuli, making quantitative comparison difficult. To remedy this, we studied MST using a family of random dot heading stimuli that we have previously used in the study of VIP. These stimuli simulate observer translation through a three-dimensional cloud of points, and a range of forward headings was presented both with and without horizontal smooth pursuit eye movements. We found that MST neurons, like VIP neurons, respond robustly to these stimuli and partially compensate for the presence of pursuit. Quantitative comparison of the responses revealed no substantial difference between the heading responses of MST and VIP neurons or in their degree of pursuit tolerance.

INTRODUCTION

The medial superior temporal area of macaque extrastriate cortex (MST) is perfectly suited for the analysis and representation of the wide-field image motions that arise from observer motion through the environment. Neurons in MST have large receptive fields (RFs) selective for different components of optic flow present in such scenes (Duffy and Wurtz 1991; Graziano et al. 1994; Tanaka et al. 1986). Additionally, they are tuned for the direction of visually simulated trajectories in a variety of visual environments (Gu et al. 2006; Page and Duffy 2003; Paolini et al. 2000). Last, MST cells explicitly represent the direction of ongoing smooth pursuit eye movements (Komatsu and Wurtz 1988).

The direct involvement of MST neurons in heading perception has been suggested by several experiments. First, MST neurons are sufficient for the discrimination of heading based on either visual or vestibular cues (Gu et al. 2007) in a two-alternative task. Second, in this task, their responses covary with perceptual decisions on a trial-by-trial basis (Gu et al. 2007). Most importantly, electrical microstimulation biases perceptual judgments of heading in a two-alternative task (Britten 1996).

The ventral intraparietal area (VIP) in the posterior parietal cortex is functionally distinct from MST in many ways. It is a multimodal area, showing visual (Bremner et al. 2002; Colby et al. 1993; Schaafmsa and Duyens 1996), vestibular (Schlack et al. 2002), somatosensory (Duhamel et al. 1998; Schlack et al. 2005), auditory (Schlack et al. 2005), and smooth pursuit-related responses (Schlack et al. 2003). Visual responses in VIP are often direction selective (Colby et al. 1993) and also selective for complex, space-varying motion stimuli (Schaafsma and Duyens 1996). Interestingly, some VIP neurons are unique in showing visual RFs in a head-centered, rather than retinotopic, frame of reference (Duhamel et al. 1997; Schlack et al. 2005). Microstimulation in VIP can produce stereotyped movements of the head and upper body, similar to defensive movements (Graziano et al. 2005). Together with the depth-limited RFs of VIP neurons (Colby et al. 1993), this has led to the suggestion that VIP is specialized for visual guidance of movement in near-extrapersonal space.

With such a range of shared and unique attributes in two anatomically linked cortical areas, it becomes critical to make precise and quantitative comparisons of their responses whenever possible. To this end, we have measured the heading tuning of MST neurons using a family of stimuli we have previously used to characterize VIP (Zhang et al. 2004). These stimuli simulate heading toward a three-dimensional cloud of points, and we simulated a range of headings varying horizontally in the forward direction. Additionally, in a subset of trials, we made the monkey track an independently moving visual target using smooth pursuit eye movements. Thus we are in a position to quantitatively compare MST and VIP using identical stimuli in an identical behavioral context. We chose to study a restricted range of frontal headings because we could more densely sample a more restricted range of headings, and thus more accurately detect changes within this range, and also because this is likely to be the most behaviorally relevant range of headings for terrestrial animals.

MST and VIP responses to these stimuli were virtually identical. In both areas, there were high-amplitude signals of heading with similar tuning widths. Furthermore, both areas were very similarly affected by the presence of horizontal smooth pursuit eye movements. From this, we conclude that both areas together support heading judgments from visual motion cues.

METHODS

Subjects and surgery

Two adult rhesus macaques (Macaca mulatta) were used in this study. Each monkey was prepared for experiments by surgical implantation of a head restraint post, a scleral search coil to monitor eye movement (Judge et al. 1980), and a chronic recording cylinder (Crist Instruments) placed over the superior temporal sulcus (STS). The
cylinder was placed in a parasagittal plane and angled 30° above horizontal, allowing dorsal-posterior access to MST. All surgical procedures were carried out under deep anesthesia in a dedicated primate surgical suite at the California National Primate Research Center. Monkeys had been trained to sit in a primate chair for the duration of the experiment with their heads restrained and to perform visual fixation and pursuit tasks in the presence of visual stimuli for juice rewards. All animal procedures were approved by the UC Davis Animal Care and Use Committee and fully conformed to ILAR and USDA guidelines for the care and treatment of experimental animals.

Stimuli

Visual stimuli were generated by custom software (Arthur L. Jones, Daniel J. Sperka) on a personal computer and displayed on a CRT monitor (Mitsubishi Diamond Pro 21TX), which subtended 72 horizontal and 56 vertical degrees of visual angle at a viewing distance of 28 cm. The spatial resolution of the display was 1,280 horizontal and 560 vertical pixels and the frame rate was 80 Hz. The fixation target was a 0.25° red circle presented superimposed on the other stimuli.

The primary stimulus used in this experiment simulated translation through a three-dimensional cubic cloud of points (Fig. 1). Dots were white (60 cd/m²) on a dark background (dim room illumination, <0.5 cd/m²). The simulated trajectory corresponded to an approach of 1 m/s horizontally and 56 vertical degrees of visual angle at a viewing distance of 28 cm. The spatial resolution of the display was 1,280 × 1,024 pixels and the frame rate was 80 Hz. The fixation target was a 0.25° red circle presented superimposed on the other stimuli.

The primary stimulus used in this experiment simulated translation through a three-dimensional cubic cloud of points (Fig. 1). Dots were white (60 cd/m²) on a dark background (dim room illumination, <0.5 cd/m²). The simulated trajectory corresponded to an approach of 1 m/s through a cloud of points 10 m across, initially centered 5.5 m from the observer. The total number of dots in view on the screen was ~2,000, and the dots filled the entire display on all trials. The dots were of fixed size throughout; all nonmotion cues to depth were removed. The simulated trajectory was varied horizontally, typically between 30° left and right of egocentric zero (directly ahead), uniformly spaced. We presented four types of randomly interleaved trials: fixation, two horizontal pursuit directions (10°/s), and pursuit-only trials where no visual stimulus was present. Each trial began with the appearance of the fixation point. After the monkey fixated, the stationary dots would appear; 250 ms later, the stimulus dots moved for 1 s. On pursuit trials, the pursuit target would begin to move 175 ms before the start of simulated self-motion to allow for smooth pursuit initiation. The monkey was required to keep its eye within a 1.5° square window around the fixation point on fixation-only trials; this window was extended horizontally to 2.5–3.5° on pursuit trials. The display was viewed binocularly, and vergence was not monitored. However, because all stimuli were at the same depth, it is likely that the animal maintained approximately constant vergence throughout each experiment. Successfully completed trials were rewarded with a drop of juice; trials where fixation was not maintained were discarded from analysis.

Electrophysiological recording

Prior to quantitative data collection, each cylinder was extensively mapped using both physiological and anatomical landmarks to localize area MST. These landmarks included recording depth, gray/white matter transitions, RF size, weak or absent retinotopy, and neuronal response properties (Saito et al. 1986; Van Essen et al. 1981). Based on the locations of our recordings relative to area MT on the posterior bank of the STS and on their response properties, we are confident that the vast majority of our sample came from the dorsal subdivision of MST, MSTd.

Neuronal activity was recorded using tungsten microelectrodes (Fred Haer) introduced through a stainless steel guide tube held in place by a plastic grid attached to the inside of the recording cylinder (Crist et al. 1988). Electrode signals were amplified and filtered, and single units were isolated with a window discriminator (Bak Electronics). Action potentials were converted to TTL pulses, the times of occurrence of which were recorded by the experimental control software (REX) (Hays et al. 1982) with a precision of 1 ms.

Once a neuron was isolated, we established the RF location and size using handheld moving bar stimuli or computer-generated moving dot patches. This was done quite carefully, using the most effective stimuli. In early experiments, the two authors independently mapped the RF, and these estimates were always close. We attempted to place the center of the heading stimuli over the most sensitive region (“hotspot”) of the RF by altering the monkey’s position of fixation. For many of our cells with large, eccentric RFs, we were forced to compromise on this ideal geometry because of the limited dimensions of the display. Cells for which we could not achieve substantial stimulus coverage of the RF were discarded.

Data analysis

Spike counts were taken from the 1-s duration of simulated self-motion, offset by 75 ms to account for visual latency. The spike counts included all trials where the animal maintained fixation; no filtering for small corrective saccades was used. Inspection of the data showed these to be rare with no clear relationship to either pursuit or eccentricity.

Neurons in both areas could be either band-pass or sigmoidally tuned across the range of frontal headings we studied. We recognize that all cells are indeed band-pass (Gu et al. 2006) if tested through a full range of headings, but we needed to capture changes in the range that we studied. We therefore needed an analysis method that would allow fair estimates of tuning for both classes of cells in both areas with quantitatively comparable parameters. We chose two related functions—probit functions (1) for the sigmoidally tuned cells and Gaussian functions (2) for the band-pass tuned cells. These functions were chosen for two reasons. They are the ones used in the Zhang et al. (2004) paper, but more importantly, the nature of the functions allows shifts of position to be independent of changes in slope. The obvious alternative (fitting all data with Gaussian functions) lacks this important benefit. It is also important to note that while the interpretations of the parameters of the two might be different with respect to the neural code for heading, changes in the parameters of the RF from pursuit may be treated the same for both. The functions we used were as follows

\[ R = A \times \left( \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{h} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right) + b \]  
\[ R = A \times e^{-\frac{(h-\mu)^2}{2\sigma^2}} + b \]

where \( R \) is neuronal response, \( A \) the amplitude, \( h \) the heading angle, \( \mu \) the mean, \( \sigma \) the width, and \( b \) the baseline. All fits were performed by the Matlab iterative fitter, fminsearch, minimizing \( \chi^2 \) error. All cells were tested for whether either of these functions accounted for significantly more of the variance than a single value; 29 cells were discarded because they failed this test. The \( \sigma \) parameters of these two
functions differ in interpretation because of their different symmetries. To solve this problem, we incorporated a scale factor to convert \( \phi \) to a single parameter we call tuning width, which describes the amount of heading that spans 50% of the dynamic range of the cell surrounding the half-maximal point (flank of the Gaussian, mean of the probit). This scale factor was 0.637 for the probit function and 0.418 for the Gaussian. This is related to the conventional half-width at half-height used in many studies. It will be slightly smaller because we estimate the range of headings that will span half the dynamic range at the functions’ steepest points, and the half-width at half-height estimates this from the peak of the tuning function. All quantitative analysis was performed on cells where the fit to one or the other of these functions was significantly better than a fit to the mean (likelihood ratio test, \( P < 0.05 \)). Comparing, for each cell, the fits from the two functions is assisted by each having the same number of parameters; we could therefore compare their fit quality fairly and without additional assumptions. For all further analysis, we used whichever function provided the best account of the data (lower \( \chi^2 \) error) under fixation; the same function was then used for all pursuit conditions. The \( \chi^2 \) values and percentage of variance accounted for by each fit type are compared for each cell in Supplemental Fig. S1.1.

**RESULTS**

We studied 105 single MST neurons in two hemispheres of two rhesus monkeys. Of these, 72 were held sufficiently long and passed our criterion for being adequately tuned. All neurons we studied had large RFs; most incorporated the fovea and often more than a quadrant of the visual field. These cells responded more strongly to large field motion than to moving bars and were selective for direction, typical of neurons in the dorsal segment of area MST (Komatsu and Wurtz 1988; Tanaka and Saito 1989). For the quantitative comparison with area VIP, we used the sample previously described (Zhang et al. 2004) with the exception that two cells with extreme \( Z > 4 \) values for tuning width were rejected, leaving 82 VIP cells.

**Heading selectivity**

As reported in other accounts (Gu et al. 2006), MST neurons are well tuned for heading when using three-dimensional (3D) optic flow stimuli. Figure 2 shows three example cells typical of our results. An additional 15 cells are shown in Supplemental Fig. S2. Cells were either sigmoidally tuned across frontal headings (Fig. 2, A and C) or else band-pass tuned (B) with a peak in the range of frontal headings we sampled. To estimate the population characteristics, we analyzed the sample distributions of mean, width and amplitude parameters from the best-fit functions (Fig. 3). Cells were tuned to a range of preferred headings, although the midpoints (half-maximal responses) of sigmoidally tuned cells tended to cluster near dead ahead (Fig. 3A). Cells showed fairly robust responses on average with robust response amplitudes and narrow tuning widths. These values are broadly consistent with previous reports, reassuring us that we were recording from a typical MST sample.

**Influence of pursuit**

Because we have previously described VIP neurons under these conditions, we were interested in how stable MST responses were under pursuit. On randomly interleaved trials, the monkeys pursued an independently moving red target superimposed on the image of the random dot cloud. In Fig. 4, we illustrate the influence of pursuit on the responses of the same three example cells shown in Fig. 2. We see that there is...
remarkably little effect of pursuit on the positions and shapes of the tuning functions. To quantify this impression, we explored the changes in the three parameters resulting from the fit functions. To do this, we independently fit functions of the same type to each pursuit condition and determined the changes resulting from the presence of pursuit.

The types of changes we measured are illustrated in the hypothetical examples in Fig. 5. Figure 5A illustrates a horizontal shift in the tuning due to pursuit. The mean shift is measured as \( \mu \) (mean parameter) of the best-fit function under each pursuit condition minus the \( \mu \) under the fixation condition. If the midpoint is stable under pursuit, the mean shift will be near zero, and if the midpoint changes under pursuit, then the mean shift will become significantly different from zero in either the positive or negative direction. Another possible influence on the tuning due to pursuit is a change in the tuning width. Figure 5B illustrates an example where the midpoints of neuronal tuning are stable but the neuron becomes less selective. Figure 5C illustrates an example where the midpoint and tuning width remain relatively stable but the gain or amplitude significantly increases. These three possible changes are of course not mutually exclusive and could in principle occur in any combination. Baseline shifts were also examined, but were not observed.

Figure 6 illustrates how the midpoints of neuronal tuning are influenced by pursuit in MST and in VIP. Tuning tended to remain fairly stable in both areas, but there was a modest shift to the right under left pursuit, and to the left under right pursuit. This slight shift was statistically significant in both areas for left pursuit, and in neither area for right pursuit because of greater scatter (see Table 1 for all statistical results). The sign of this shift was consistent with what has been observed in both areas in previous work: modest under-compensation for pursuit (the sign of the compensation was stated erroneously in our previous report). Using depth-containing stimuli such as ours, it is not possible to define a single retinal shift value that would be expected without compensation, since the amount of retinal shift depends on the simulated depth corresponding to any image vector. However, in our stimuli, the average shift across all depths was \( \sim 40^\circ \), and the median shifts in our data were 3–5°. Therefore it can be safely concluded that in both areas, the representation is close to completely compensated for pursuit [which we previously termed “head-centered” (Zhang et al. 2004), with respect to pursuit]. In MST, there is also a modest yet significant relationship between center tuning and the degree of shift. Cells with more central tuning tended to be slightly more shifted under pursuit than were cells with more eccentric preferences (left pursuit: \( r = 0.14, P < 0.0014 \); right pursuit: \( r = 0.07, P < 0.023 \)). There might be technical contributions to this result, which we consider in the Discussion.

Response amplitudes were also very similar in both areas although there was a slight tendency for MST cells to have higher amplitudes. This difference was not significant in a Bonferroni-corrected test for equality of median. The influence of pursuit is represented in Fig. 7. The main features of these scatterplots in both areas are high correlation between values estimated from the fixation and pursuit trials, and the tendency for the points to lie along the diagonal line that depicts identity. The mean was not significantly shifted away from the diagonal (Table 1).

The one statistically robust difference between the tunings in MST and VIP was in tuning width: MST tuning widths were significantly larger than those in VIP (\( P < 0.001 \), Wilcoxon

![Diagram](image-url)
test). This difference suggests that if anything, the representation in VIP is more precise than in MST when tested with stimuli such as ours. The influence of pursuit on the tuning widths in MST and VIP is shown in Fig. 8. There was a modest systematic increase in width in both areas of comparable magnitude. This is consistent with known effects of pursuit on the perception of heading in monkeys (Britten and Van Wezel 2002); increased tuning width should lead to increase in threshold. In three of four comparisons (excepting VIP for right pursuit), this increase was significant or marginally significant (Table 1).

We were naturally curious about whether the compensation that we observed was due to explicit signals of pursuit in either area. We explored the relationship between each of the metrics we examined and responses to pursuit in the dark (Supplemental Figs. S3 and S4 for MST and VIP, respectively). While substantial pursuit responses were observed in both areas, these were not predictive of any measured tuning changes under pursuit.

**DISCUSSION**

This experiment produced two main findings. First, in agreement with previous observations, we find that MST neurons represent heading stimuli well, and this representation is stable; however, we found that VIP neurons are also capable of representing heading stimuli, albeit with less precision. This suggests that the representation in VIP is more precise than in MST when tested with stimuli such as ours. The influence of pursuit on the tuning widths in MST and VIP is shown in Fig. 8. There was a modest systematic increase in width in both areas of comparable magnitude. This is consistent with known effects of pursuit on the perception of heading in monkeys (Britten and Van Wezel 2002); increased tuning width should lead to increase in threshold. In three of four comparisons (excepting VIP for right pursuit), this increase was significant or marginally significant (Table 1).

**TABLE 1. Statistical testing of tuning parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MST</th>
<th>VIP</th>
<th>MST = VIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center</td>
<td>1.38</td>
<td>0.93</td>
<td>0.37</td>
</tr>
<tr>
<td>Amplitude</td>
<td>27.39</td>
<td>21.59</td>
<td>0.01</td>
</tr>
<tr>
<td>Tuning width</td>
<td>13.10</td>
<td>4.19</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Center shift - l.p.</td>
<td>3.70</td>
<td>4.64</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Center shift - r.p.</td>
<td>-3.00</td>
<td>-2.77</td>
<td>0.49</td>
</tr>
<tr>
<td>Ampl. change - l.p.</td>
<td>-1.04</td>
<td>-0.31</td>
<td>0.86</td>
</tr>
<tr>
<td>Ampl. change - r.p.</td>
<td>-1.79</td>
<td>-1.17</td>
<td>0.25</td>
</tr>
<tr>
<td>Width change - l.p.</td>
<td>1.14</td>
<td>0.02</td>
<td>0.94</td>
</tr>
<tr>
<td>Width change - r.p.</td>
<td>1.12</td>
<td>1.02</td>
<td>0.003</td>
</tr>
<tr>
<td>Bl. change - l.p.</td>
<td>-0.08</td>
<td>0.00</td>
<td>0.62</td>
</tr>
<tr>
<td>Bl. change - r.p.</td>
<td>0.00</td>
<td>1.54</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Results of nonparametric statistical tests. The single area hypotheses under test were whether the sample median differed from zero (Wilcoxon signed-rank test). In the last column we test the hypotheses that the sample medians are the same in MST and VIP (Wilcoxon rank sum test). Bold-print entries are those that remain significant after Bonferroni correction for multiple tests (N=27, P < 0.002). Abbreviations: l.p., left pursuit; r.p., right pursuit; ampl., amplitude; bl., baseline.
in the face of smooth pursuit eye movements. Second, we find that most of the tuning properties of MST and VIP are quite similar with the exception of tuning width, where VIP is more tightly tuned than is MST.

Technical issues

Several decisions were made in the design of this experiment with potential consequences for the resulting data. All were driven by the desire to keep conditions identical to the experiment of Zhang et al. (2004). One fundamental decision was to only explore frontal headings. While MST neurons represent the full 360° range of headings in both azimuth and elevation (Gu et al. 2006), we are only able to describe accurately those that significantly modulate within a 60°, horizontal range of forward headings. These are presumably the most relevant signals for normal locomotion for terrestrial vertebrates, even if they are a select subpopulation of the neurons in the structure. These are also the neurons mostly likely to contribute to performance on the most commonly studied two-alternative heading discrimination task (Britten and Van Wezel 1998, 2002; Gu et al. 2008; Zhang and Britten 2010). We recognize that we cannot conclude from our data that all ranges of headings would be as similar in MT and VIP as the parts we describe, but it would be a very strong prediction. No work has to our knowledge shown any special properties of neurons preferring forward headings in either area.

One potential consequence of restricting the range of headings sampled is more technical; it might introduce biases into our curve-fitting and parameter estimation. While this is possible, we consider it unlikely. Our maximum-likelihood fitting method weights all the headings in our data equally, and the fit parameters (especially the mean of the Gaussian) were not constrained to remain within the sampled range. Inspection of the fits and of the residuals showed no systematic departures for any group of neurons in either area. A result that is suspicious in this regard is the relationship between the amount of shift and the center tuning in MST (Fig. 6). If our fit parameters were less stable for more eccentrically tuned cells, such a result might occur. We believe this to be unlikely for several reasons. Any such instability should in principle be a source of variance but not bias. Independent fits were made to both the pursuit and fixation data, and whatever biases might apply would presumably apply to both, and thus be removed by the subtraction that produced the shift. Also any uncertainty should produce increased variance because the variance of the shift would be equal to the sum of the variances of each center parameter. Analysis of the scatter of the shift against center tuning showed no relationship (regression, \( P > 0.2 \) for both pursuit directions). While no previous report has described a relationship between pursuit compensation and center tuning, this question might deserve further study. It would ideally be approached in a design intended to sample a wider range of headings than ours was.

One important aspect of our approach was that we moved the animal’s fixation point to maximize overlap between stimulus and each cell’s RF. This is a bit unusual, and we need to consider the consequences for the data. The choice allowed us to study carefully a greater fraction of the cells we recorded. If we kept the fixation constant, many cells would have responded much more poorly and would have been discarded before quantitative analysis. What we instead get is some kind of upper envelope of cell response amplitude and tuning quality. This reduces the extent with which our data can be rigorously compared with other labs’ data; but critically, because VIP was studied in exactly this way, it was necessary to do the same for MST to enable the comparison that is the main focus of this paper. Responses in both VIP and MST are both affected by gaze direction and smooth pursuit (Duhamel et al. 1998; Fetsch et al. 2007; Newsome et al. 1988; Schlack et al. 2003; Squatrito and Maioli 1997). We will consider gaze and pursuit effects individually.

MST responses are modulated by eye position (Fetsch et al. 2007; Squatrito and Maioli 1997), whereas VIP neurons often show shifts of RF location that (in the extreme) fully compensate for the change in gaze, producing a cell with a head-centered RF (Duhamel et al. 1998; Schlack et al. 2005). The range of gaze angles we employed was small compared with what was used in these studies (± 20° from straight ahead, both vertically and horizontally), and for each cell, the gaze direction never changed >5° from its average. We analyzed subsets of the data from the middle 250 ms of each response epoch, reducing the gaze changes to 1.25°, and no differences in the results were seen (data not shown). For both of these reasons, we are doubtful that gaze dependence influenced our analysis in any meaningful way.

Another choice was the form of the functions we used to describe the data. The data in both VIP and MST required both a sigmoid and a peaked function; this was merely a secondary consequence of our bounded stimulus range. We know from first principles that all tuning functions must be peaked if sampled widely enough because the azimuth axis is circular. We therefore examined the consequence of using Gaussians to describe all the data sets—the flank of the Gaussian would form a sigmoid across the range of headings we sampled. When compared against the probit fits, these systematically gave greater \( \chi^2 \) errors because of the asymmetry of the sigmoid. The data were clearly more symmetrical with approximately equal curvatures on either side of the point of maximum slope. Another unfortunate feature of using Gaussian functions for the sigmoid data was apparent from inspection of how the fit parameters were affected by pursuit. Frequently, changes in amplitude would lead to spurious changes in the center point and/or the sigma parameter, and these were frequently very large and unpredictable. The probit gave much more sensible changes because the midpoint parameter was located near the steeply changing part of the data and was more independent from the width and amplitude parameters.

Pursuit interactions in MST

Signals of ongoing smooth pursuit are common in the upper levels of the motion system, but the function of these signals is not at all clear. Early studies revealed the presence of pursuit signals in MST and that these were not merely additive and independent (Komatsu and Wurtz 1988; Thier and Erickson 1992).

Since that time, a large number of papers have explored the question of compensation of tuning in MST in the presence of pursuit. The principal question here is whether MST can distinguish external object motion from reafferent motion from eye movements. While different labs have used quite different
stimuli, the main questions are quite closely related. In one set of experiments, the stimuli are spatially uniform and speed is varied (Chukoske and Movshon 2009; Inaba et al. 2007). This situation is relatively straightforward to interpret because predictions can be formulated for any amount of shift in the speed tuning curves, from purely retinal to completely compensated. The latter would suggest the complete discounting of reafferent motion, allowing the representation to be read out as external object motion. The results are not nearly so clean-cut, however, and most cells in MST exhibit something in between the two extremes. Additionally, changes in gain or tuning curve shape are not uncommon. However, setting such complexities aside, the Inaba et al. report (Inaba et al. 2007) appears to document systematically greater compensation than does the Chukoske et al. paper (Chukoske and Movshon 2009). While these experiments were fairly similar, one striking difference was in the spatial extent of the stimuli: Inaba et al. used full-screen stimuli, whereas Chukoske et al. used stimuli restricted to the “hotspot” of the RF.

Many other experiments have followed the suggestion that MST contains a representation of self-motion direction. These used more complex, space-varying stimuli simulating observer translation, typically toward a 2D frontoparallel plane of dots or a 3D cloud of dots (as we did in this study). The Duffy lab has reported quite stable representations of heading tuning, when heading is varied circularly around dead ahead (Page and Duffy 1999). This stability is further improved by the addition of multiple depth planes to the stimulus (Upadhay et al. 2000). However, in these experiments, responses were not compared against a retinal prediction, so the exact degree of stability is uncertain. Work from the Andersen and Bremmer labs has specifically targeted this question (Bradley et al. 1996; Bremmer et al. 2009; Shenoy et al. 1999). While the details of stimuli and analyses differ, the consensus finding from these experiments is that pursuit compensation is partial. In the Andersen lab studies (Bradley et al. 1996; Shenoy et al. 1999, 2002), depth was removed from the stimuli to allow quantitative comparison of compensation of real and simulated pursuit. In the “real pursuit” condition most comparable to ours, responses appeared to be more highly compensated in real pursuit, compared with simulated pursuit. This strongly supports the use of extraretinal signals in pursuit compensation. The presence of simulated depth in our experiments makes direct comparison with these studies difficult for two reasons. The first is technical. Because of limitations in our custom software, we could not generate simulated-pursuit stimuli. More importantly, the retinal shift in our experiments was not a unitary value. On the near side of the dot cloud, where retinal velocities were highest, the pursuit-induced shift was negligible. For the farthest dots, it was ~80°. Our observed values of 3–5° were small by comparison with the average (40°), but the average might not be the appropriate benchmark. For instance, one could imagine that a purely visual mechanism, which weighted the highest-velocity dots in the image most heavily, could achieve pursuit invariance straightforwardly. While it is evident from other experiments, most notably the Andersen lab work cited in the preceding text, that extraretinal signals can support pursuit tolerance, it seems likely that visual mechanisms might play a role when they can, in the presence of image depth. This was not the primary goal of this experiment, but the present work highlights the importance of doing a directed experiment on this critical mechanistic question in both MST and VIP.

There is greater similarity between the speed- and heading-tuning experiments than might at first meet the eye. In a heading stimulus, many speeds and directions are present in the display and usually within the RF of the MST cell. Therefore changes in the speed tuning of the neuron to single speeds could possibly underlie the stability of tuning to these more complicated stimuli. Indeed such effects must occur for some MST neurons: ones with a high degree of translation-invariance in their tuning to complex motion patterns (Duffy and Wurtz 1991; Graziano et al. 1990; Saito et al. 1986). These would seem ill-suited to representing heading direction because the dominant change that occurs with changes of heading direction is a change in the retinal location of the expanding optic flow pattern. Translation invariance would then appear to defeat heading tuning. However, because the distribution of speeds in the RF will change as well with changes in heading, then speed tuning could contribute to both the heading tuning of such cells and their pursuit tolerance. Indeed this idea lies at the heart of some models of heading (Perrone and Stone 1998). All of these considerations together, though, suggest that to more completely understand visual contributions to pursuit compensation at a mechanistic level, future experiments will need to focus on both how multiple speeds interact in MST cells’ RFs.

**MST and VIP**

The primary goal of this work was the direct and quantitative comparison of heading responses in MST and VIP using exactly the same stimuli. We found remarkable similarity of the heading representations in the two areas. The only substantial difference we uncovered was that VIP neurons possessed somewhat tighter tuning (narrower width) than did MST neurons (see Table 1).

The literature also shows a great degree of similarity between MST and VIP. Both areas receive direct feed-forward input from MT (Lewis and Van Essen 2000; Maunsell and Van Essen 1983; Ungerleider and Desimone 1986), and both contain a large majority of strongly directionally selective neurons (Colby et al. 1993; Duffy and Wurtz 1991; Graziano et al. 1990; Schaafsmat and Duysens 1996; Schaafsmat et al. 1997; Tanaka et al. 1986). Both contain qualitatively similar signals of ongoing pursuit eye movements, thought to be extra-retinal in origin (Bremmer et al. 1996; Komatsu and Wurtz 1988; Schlack et al. 2003), as well as vestibular signals (Gu et al. 2006; Schlack et al. 2002). While differences in the details of the stimuli make quantitative comparison difficult, it seems likely that formal study of the comparison would make the similarities more profound.

However, there are some striking differences between the two areas as well. Many VIP neurons respond to somatosensory stimulation (Colby et al. 1993; Schlack et al. 2005), but MST neurons apparently do not. While imaging studies of human MST have revealed somatosensory responses (Beauchamp et al. 2007, 2009), these have not been seen when it was specifically examined in a single-neuron study in anesthetized monkeys (Hikosaka et al. 1988). VIP also shows auditory responses, which often show
strikingly delimited spatial receptive fields (Schlack et al. 2005), yet auditory responses have not been reported in area MST.

So what does one conclude from such a mixture of results? The immediate conclusion is that neither area is specialized for any single perceptual or behavioral role. But this begs the question of what more positive conclusions can be drawn. Clearly, the answer lies in mapping these areas to particular tasks that can be studied in the laboratory, and heading perception is probably the most developed such task, on both theoretical and empirical grounds. In this context, based on the present results and the other related studies cited above, MST and VIP appear to contribute largely equivalent information concerning heading based on visual cues. Both areas effectively combat the problem of reafference and represent heading with eye rotations largely removed. In a previous publication, we referred to this as “head-centered coordinates” (Zhang et al. 2004), but we wish to make clear this is not in the sense of Cartesian spatial coordinates but only with respect to eye and head velocity. In any case, the similarity of the two areas suggests a testable hypothesis that each contributes equally to the perception of heading. This hypothesis can be tested with a variety of tools, including dual lesions, simultaneous recordings, and dual microstimulation experiments. Another approach is to design heading tasks based on cues not held in common between MST and VIP (e.g., auditory or somatosensory) to find the limits of the equal-contribution hypothesis. Undoubtedly, in natural behavior, the suite of cues used by primates to determine heading is large, and the weighting of these cues is clearly flexible (Gu et al. 2008). Recent theory suggests that near-optimal cue integration arises naturally from the statistical properties of multimodal neuronal responses (Ma et al. 2006), but this needs explicit testing in the domain of heading perception.

ACKNOWLEDGMENTS

The authors thank H. Engelhardt for care and training of the monkeys as well as D. Sperka and A. Jones for custom software and support. We also thank C. McCool, T. Zhang, and S. Egger for carefully reading earlier versions of the paper.

GRANTS

This work was supported by National Eye Institute Grant EY-10562 to K. H. Britten. J. B. Maciokas was partially supported by T32 EY-015387 (principal investigator: J. S. Werner). Additional support came from Vision Core Center Grant EY-12576 (primary investigator: L. M. Chalupa).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES


J Neurophysiol • VOL 104 • JULY 2010 • www.jn.org


