Differential Neuronal Representation of Spatial Attention Dependent on Relative Target Locations during Multiple Object Tracking

Ayano Matsushima and Masaki Tanaka
Department of Physiology, Hokkaido University School of Medicine, Sapporo 060-8638, Japan

Humans can simultaneously track multiple moving objects with attention. The number of objects that can be tracked is known to be larger when visual stimuli are presented bilaterally rather than presented unilaterally. To elucidate the underlying neuronal mechanism, we trained monkeys to covertly track a single or multiple objects. We found that neurons in the lateral prefrontal cortex exhibited greater activity for the target passing through the receptive field (RF) than for distractors. During multiple-object tracking, response enhancement for one target presented in the RF was stronger when the other target was located in the opposite than the same visual hemifield. Because the neuronal modulation did not differ depending on relative target locations with respect to upper and lower visual hemifields, the distance between the targets does not explain the results. We propose that inherent, anatomical separation of visual processing for contralateral and ipsilateral visual fields might constrain cognitive capacity.

Key words: attentional capacity; bilateral advantage; MOT; prefrontal cortex; primate; single-unit recording

Introduction

Along the visual processing pathways, the left visual field is represented in the right hemisphere and the right visual field is represented in the left hemisphere. This laterality is especially evident in the early stages of visual processing, but is relaxed for higher order processing in the association cortices where information from both visual hemifields are integrated (Desimone et al., 1993; Rainer et al., 1998). This integration process must be mediated via the commissural connections, which unify the representation of the entire visual field.

Deviation from the above principle may occur in split-brain patients (Gazzaniga, 1970). Although disconnected cerebral cortices can exchange crude visual information, they cannot communicate details about features, locations (Holtzman, 1984), or time (Corbalis et al., 1998). Conversely, split-brain patients are less annoyed by interference across the midline when stimuli are presented bilaterally, exhibiting superior performances to normal subjects in the Stroop task (David, 1992), dual task paradigm (Holtzman and Gazzaniga, 1985), and visual search (Luck et al., 1989). Thus, the compromised interhemispheric integration can hinder or benefit the performance depending on whether the task requires integration or separation of information from different visual hemifields.

Even in normal subjects, the integration across hemifields has been shown to be incomplete. In the tasks requiring direct comparison of multiple objects (Banich and Belger, 1990; Sergent, 1990) or integration of visual information across time (Dimond et al., 1972), performance is impaired when the stimuli are divided into the left and right visual fields. However, in the tasks requiring parallel processing of individual objects, performance is improved when the stimuli are presented across hemifields compared with within a hemifield (Sereno and Kosslyn, 1991; Awh and Pashler, 2000; Scafà et al., 2007). These costs and benefits of bilateral display are most evident when the task is rather demanding (Merola and Liederman, 1990; Kraft et al., 2005) and requires spatial information processing (Delvenne, 2005).

At the neuronal level, evidence has recently accumulated for parallel processing of two visual hemifields. When two visual stimuli were presented bilaterally, prefrontal activities were initially dominated by the contralateral stimulus, and thereafter evolved to represent the task-relevant object (Kadohisa et al., 2013). In addition, when monkeys remembered the locations (Matsushima and Tanaka, 2014b) or identities (Buschman et al., 2011) of multiple objects, neuronal information for a given object was not, if any, degraded when all other objects were in the opposite hemifield. These results suggest that neuronal processing for visual discrimination and working memory is more efficient for bilateral than unilateral stimuli. However, it remains elusive whether the attentional allocation toward the targets also depends on their relative locations across the hemifields, while keeping it away from the other, task-irrelevant stimuli.

To address this issue, we examined neuronal activities of prefrontal neurons during monkeys attentively tracked two targets...
in the presence of distractors. We found stronger response to the targets presented bilaterally than those presented unilaterally, which might explain the different tracking capacity of humans under the two conditions (Alvarez and Cavanagh, 2005; Hudson et al., 2012).

Materials and Methods

Experiments were conducted on two Japanese macaques (Macaca fuscata, 6–7 kg females). All animal protocols were approved by the Animal Care and Use Committee of Hokkaido University. The procedure of animal preparation is described previously in detail (Tanaka, 2005). Briefly, a pair of head holders was implanted on the skull and a coil of stainless steel wire was implanted under the conjunctiva under isoflurane anesthesia. During training and experimental sessions, the monkey’s head was secured to the primate chair, and eye position was recorded using the search coil technique (MEL-25; Enzanshi Kogyo). After training, a recording cylinder was installed over a small craniotomy. Animals received analgesia after each surgery. The monkeys’ water intake was controlled daily to motivate them to perform the tasks.

Visual stimuli and behavioral paradigms. Experiments were controlled by a Windows-based real-time data acquisition system (TEMPO; Reflective Computing). Visual stimuli were presented on a 24 inch CRT monitor (GDM-FW900, refresh rate 60 Hz; Sony) positioned 38 cm from the eyes (subtending 64° × 44° of visual angle), located in a darkened booth. All visual stimuli were presented within a 36° square contour that was visible throughout the experiments. Each trial began with the appearance of a fixation point (FP; 0.5° red square) at the screen center, and ended with a juice reward for correct performance.

In the multiple-object tracking (MOT) task (Fig. 1A, left), four visual stimuli (1.5° white circles) were presented at 10° eccentricity during central fixation. Their initial locations and motion directions were defined in polar coordinates, and angles were chosen randomly from 0° to 350° (10° increments, measured from rightward) for each object. The separation of polar angles between any two stimuli was set to be ≥50°, for both the initial locations and motion directions. Two objects were cued as targets by briefly changing their color (red, 700 ms), whereas the objects remained white and served as distractors. Except for this cue period, all objects were identical in color and shape. Following a 200 ms delay, four objects moved along straight paths at 20°/s in different directions without visible trails and bounced against the sides of the contour (motion period: 2500 ms for Monkey J, 2000 ms for Monkey L). Each object could move into the other quadrants than the initial location. Monkeys were required to keep their eyes within 5° of the FP throughout the motion and following delay (500 ms) periods, and make two sequential saccades to the targets. To obtain reward, the animals needed to make the initial saccade within 400 ms of the FP offset and to redirect their eyes to the other target in 400 ms thereafter. When any object was located within 7° from the FP or the distance between any two objects was <6°, the motion interval was extended until neither condition was fulfilled. We also introduced the single-object tracking (SOT) task (Fig. 1A, right) as a control. This task was identical to the MOT task except that only one object was cued as a target.

Once the animals performed correctly in a MOT trial, the object trajectories in the subsequent seven MOT trials and eight SOT trials were automatically determined to balance the object trajectories across trials (Fig. 2A, bottom). In three MOT trials, the object trajectories were identical to the initial successful trial but the assignment of object identities was rotated one by one. In the other four MOT trials, the object trajectories were point symmetry of those in the above-described MOT trials with respect to the central FP. In the eight SOT trials, the object trajectories were exactly the same as the MOT trials. Thus, once monkeys performed a MOT trial correctly, the object trajectories in the following 15 trials were determined automatically, and were presented in a pseudorandom order.

Physiological procedures. Neuronal activity was recorded through tungsten electrodes (Alpha Omega Engineering) lowered into the prefrontal cortex (PFC) through a 23 gauge guide tube using a micromanipulator (MO-97S; Narishige). Signals were amplified (Model 1800; A-M Systems), filtered, and monitored online using oscilloscopes and an audio device. Waveforms of single neurons were isolated using a real-time spike sorter with template-matching algorithms (MSD or ASD; Alpha Omega Engineering). Occurrences of action potentials were time stamped and saved in files with the data for eye movements and stimulus locations (sampling rate, 1 kHz).

Verification of recording sites. A postmortem examination of recording sites was performed in all monkeys (Fig. 1B). At the end of the experiments, the animals were deeply anesthetized with a lethal dose of sodium pentobarbital (> 50 mg/kg, intraperitoneally), and several landmark pins were penetrated at known coordinates. The animals were then perfused with 0.1 M phosphate buffer followed by 3.5% formalin. The brain was removed, blocked, and fixed with the same solution overnight. Location of each task-related neuron was reconstructed according to the coordinates of electrode tip relative to the landmark pins.

Data analysis. Data were analyzed offline using MATLAB (MathWorks). We considered only neurons that were tested for >80 trials (n = 137). To assess the firing modulation depending on the target and distractor locations during the motion period, the instantaneous firing rate was calculated as a function of each object location on the retina (i.e., location on the screen minus eye position in the orbit). Initially, each object location 60 ms before the occurrence of action potentials was plotted separately, with a 0.5° resolution. Then, the number of action potentials was divided by the duration of each object presentation to obtain the firing rate for each pixel (0.5° square). Data were filtered using a 2-D Gaussian kernel (σ = 2.5°, 20° square), and presented as a color-coded map (“response modulation map”).

For individual neurons, we defined the RF in the response modulation map for target locations in the SOT trials as the region with 477 adjacent pixels (10% of whole map) that gave the maximal mean firing rate. The RF center was defined as the center of gravity weighted by the firing rate for each of the 477 pixels. In Figure 3, we examined whether the response to one target in the RF was modulated depending on whether the other target appeared in the different (“Across” condition) or same (“Within” condition) visual hemifield. To reliably distinguish these conditions, we only considered neurons with RF centers that were ≥6° eccentricity and located ≥3° away from the cardinal meridians (83/137, 61%).
To estimate the null response modulations observed by chance, we shuffled the correspondence between spike data and object trajectories in individual neurons, and computed a “sham” response modulation map (Matsushima and Tanaka, 2012). We repeated this shuffling procedure 1000 times to obtain the distribution of variance of firing rates across the field, expected by chance. Neurons exhibiting a significant firing modulation in the SOT trials (paired \( t \) test, \( p < 0.05 \)) were included for further analyses.

For the population analysis shown in Figures 3 and 4, the response to the targets in the RF was calculated by subtracting the mean firing rates in the whole response modulation map. To directly rule out the possibility that the Across versus Within difference stems from the different target separations, we compared neuronal responses to the two targets separated by a same distance but located within or across the visual fields. In search of the optimal target separation, we first plotted the firing rates as a function of target separations (bin width = 10°, 1° increments), regardless of the Across/Within conditions. To disregard data when both targets were within the RF, bins centered ±1.5 times of the maximum point separation within the RF were excluded from the analysis. If the optimal target separation lasted for >1000 ms in total, for both the Across and Within conditions, we compared the distributions of millisecond-by-millisecond firing rates (filtered by Gaussian with 30 ms \( \sigma \)) under the Across and Within conditions (Wilcoxon rank sum test, \( p < 0.05 \)), whereas the variances did not (Ansari-Bradley test, \( p > 0.05 \)). When the same data were divided according to the upper/lower visual hemifields, the activities were not different (median, \( p = 0.54 \); variance, \( p = 0.29 \)).

Results
We recorded from single neurons in the PFC close to the frontal eye fields while monkeys covertly tracked moving object(s) without eye movements (Fig. 1A). During recording sessions, both monkeys performed very well. Deviation of eye position from the center of the FP during covert tracking averaged 0.68° ± 0.11° (SD) and 0.73° ± 0.10° for Monkeys J and L, respectively. Both monkeys correctly chose targets in the SOT (mean ± SD, 77.5 ± 7.6% and 88.2 ± 6.0% for Monkeys J and L, respectively) and the MOT trials (64.7 ± 10.0% and 54.4 ± 7.2%), well above the chance level (25% for SOT or 8.3% for MOT trials).

Regardless of the number of objects to be tracked, PFC neurons consistently signaled the target positions during the motion period. Figure 2A shows the response modulation maps for a representative neuron. In the SOT trials (Fig. 2A, two top left panels), this neuron exhibited enhanced activity as the target passed through the upper-left RF, whereas the activity was less for the distractor presented at the same location. Since the trajectories of the target and distractor were exactly balanced across the trials (Fig. 2A, two bottom left panels), the difference in response must be attributed to their different meanings in the task. Although neuronal response to the stimulus outside of the RF ap-
peared to be somewhat larger for the distractor than for the target, it was merely a consequence of the enhanced response to the target presented in the RF, because the same spike data were sorted by the target and distractor locations, respectively. The pattern of neuronal modulation was consistent with our previous results of the Target-selective prefrontal neurons in a similar SOT task (Matsushima and Tanaka, 2012), and those reported in the extrastriate visual areas (Mitchell et al., 2007; Niebergall et al., 2011a). In the MOT trials, the same neuron again exhibited a greater response to the target than to the distractor presented in the RF (Fig. 2A, two top right panels), indicating that this neuron was involved in the representation of multiple as well as single-target locations.

We next examined whether the response to the target in the RF depended on the relative locations of two targets in the MOT trials. We divided the data into two groups, according to whether the targets were presented bilaterally or unilaterally. When the two targets were located across the left and right visual hemifields (Across condition), the neuronal response to the target was strong (Fig. 2B, left). However, when the two targets were located within either visual hemifield (Within condition), the response was reduced (Fig. 2B, middle). To quantitatively compare these two conditions, we computed the millisecond-by-millisecond firing rates (filtered by Gaussian with 30 ms σ) while the target was presented in the RF, and found that the response was statistically greater under the Across than Within conditions (Wilcoxon rank sum test, \( p < 0.0001 \); Fig. 2B, right). As a control analysis, we also divided the same data according to whether the targets were presented across or within the upper/lower visual hemifields, and found no difference between the conditions (\( p = 0.54 \); Fig. 2C).

Selective enhancement of response to the target, and the reduced target response in the Within condition were also observed across the population we studied. Among 82 neurons exhibiting significant firing modulation in the SOT trials (Methods), 80 (98%) neurons also showed a significant modulation in the MOT trials. However, the magnitude of the response to the target was slightly but significantly less in the MOT than in the SOT trials (paired \( t \) test, \( p < 0.05 \); Fig. 3A), with a robust regression slope of 0.68, (significantly different from 1, \( t \) test, \( p < 0.0001 \)). As the data were sorted according to the relative target locations, the response was greater when the two targets were presented across the left and right visual hemifields (Across condition) than when presented within either hemifield (Within condition, \( p < 0.05 \); Fig. 3B). Accordingly, the robust regression analysis yielded the slope of 0.58 that significantly differed from the unity line (\( p < 0.0001 \)). In contrast, there was no significant difference between the Across and Within conditions with regard to the upper versus lower visual hemifields (\( p = 0.45 \); Fig. 3C), and the robust regression slope did not differ from 1 (0.89, \( p > 0.1 \)). The significant difference in the whole population indicates that the target representation in the PFC was degraded when both targets were presented unilaterally, which might cause an overall reduction of firing modulation in the MOT trials compared with the SOT trials (Fig. 3A).

Different neuronal modulation for targets presented within and across hemifields during motion periods were unlikely to be related to the difference in saccade directions, because targets moved randomly before the execution of saccades. Nevertheless, because we previously showed that prefrontal activities biased the selection of an object to be tracked then targeted by a saccade at the trial end (Matsushima and Tanaka, 2014a), neuronal activities during the motion periods might control the saccade orders at the end of the MOT trials. To see whether saccade orders could account for the differential response between the Across and Within conditions, we further divided the data according to the saccade order. Although some neurons exhibited a slight difference (Fig. 4A; top versus bottom, for the neuron shown in Fig. 2), the magnitude of neuronal modulation by saccade orders was comparable between the Across and Within conditions in the population (paired \( t \) test, \( p > 0.9 \); Fig. 4B). Furthermore, the response to the target presented in the RF was consistently larger in the Across than Within condition, regardless of whether the target was selected by the first (Fig. 4C) or second (Fig. 4D) saccade. Thus, the difference in neuronal activity between the conditions could not be attributed to the serial order of targeting saccades. Although the neuronal control of saccade sequence is an interesting issue to be addressed (Histed and Miller, 2006), it appears to be far beyond the scope of this study.

One might argue that the difference between the Across and Within conditions could be merely a consequence of surround suppression geometry. However, surround suppression from distractors was unlikely to explain the results. We searched for the Across condition where a target was presented in the RF and both distractors were presented either in the same or opposite hemifield. For 13 neurons that fulfilled these criteria, neuronal activities remained unchanged regardless of the distractor locations (paired \( t \) test, \( p = 0.15 \)). Furthermore, the Across/Within difference could not be attributed to the relative distance between the targets. When the millisecond-by-millisecond firing rates of in-
Individual neurons were compared when the targets were presented within or across visual fields but separated by a same distance (see Materials and Methods), the majority of neurons (48/76, 63%) exhibited higher activity in the Across than Within conditions (t test, p < 0.05). We further confirmed that the restriction of the analysis shown in Figure 3 to these 48 neurons reproduced the main results (Fig. 3, blue dots); the robust regression slope significantly differ from 1 for the comparison between the Across and Within conditions with regard to the left/right visual hemifields (0.64; Fig. 3B, blue line; t test, p < 0.0001), but was not with regard to the upper/lower visual hemifields (0.95; Fig. 3C, blue line; t test, p > 0.1). Thus, prefrontal neurons exhibited greater response when the two targets were presented bilaterally than unilaterally, and that these modulations could not be accounted for other factors than the relative target locations across the left and right visual fields.

As an additional analysis, we examined the variability of neuronal activity to assess whether the response reliability was also altered. For example, for the neuron shown in Figure 2, the variance of response to the target in the RF was larger in the MOT than in the SOT trials (Ansari-Bradley test, p < 0.05). However, the effect of the task condition was not consistent across the population; the variance of neuronal activity in the MOT trials could be larger than (16%, 13/80), smaller than (41%, 33/80), or comparable with (43%, 34/80) that in the SOT trials. Similarly, the effect of relative target locations on the response variance was not evident for the majority of neurons (41/80, 53%, p > 0.05) including the one in Figure 2B, whereas the activity of the remaining neurons was either less (22/80, 27%) or more (16/80, 20%) variable in the Across than Within conditions. To assess the overall trend, we computed Fano factors by normalizing the variance of firing rates by their mean for individual neurons. For both comparisons between the tasks (i.e., MOT vs SOT) and between the Across/Within conditions, Fano factors were not significantly different (n = 80, paired t test, p > 0.1), indicating that the response reliability remained unchanged in the population as a whole.

**Discussion**

We examined single neuron activities in the PFC while monkeys covertly tracked a single or multiple moving target(s) among visually identical distractors. We found that the same set of neurons elevated the activity for the target presented in the RF, regardless of the number of objects to be tracked. Enhancement of the response to a target was greater when the other target was located in the opposite than the same visual hemisphere. Because the neuronal activity did not alter depending on the relative target locations with respect to the upper/lower rather than the right/ left visual fields, the distance between the targets should not account for the results. Considering the inherent, anatomical separation of contralateral and ipsilateral visual processing, our data might reflect the stronger interaction of attentional signals for multiple targets presented unilaterally than those presented bilaterally.

**Interaction of information processing within and across visual hemifields**

During multiple-object tracking, prefrontal neurons exhibited greater activity for the target in the RF when the other target was presented in the opposite than the same visual field. Different neuronal modulation between the Across and Within conditions were likely to reflect different amount of interaction in neuronal networks; signals from opposite visual fields are processed separately, whereas those from the same visual field are interacted and processed together within the same functional module.

What network configurations are responsible for such a functional module of visual processing? One possible structure is the cerebral hemisphere. Anatomically, after the optic chiasm, visual inputs from the left and right visual fields are divided into different optic tracts. The separation is precisely retained in the early visual cortices, where neurons in each hemisphere are exclusively responsive to contralateral stimuli (Bullier, 2004). Although commissural connections somewhat alleviates the hemispheric independence in higher-order cortical areas (Myers and Sperri, 1958), much stronger association fibers within the hemispheres results in the contralateral dominance of visual processing in the inferotemporal (Chelazzi et al., 1998) and prefrontal (Funahashi et al., 1989; Lennert and Martinez-Trujillo, 2013) cortices. Indeed, inactivation of the PFC has demonstrated the contralateral bias of mnemonic (Funahashi et al., 1993) or attentional (Heide and Kömpf, 1998) deficits, indicating that the functional division between hemispheres is, on the whole, preserved in the PFC. In addition, other anatomical structures, such as the cortical columns, might further dissociate visual processing in the opposite visual fields. In the PFC, commissural fibers from the contralateral counterpart are known to terminate in columns that are distinct from, and interdigitated with, those receiving association fibers from the ipsilateral parietal cortex (Goldman-Rakic and Schwartz, 1982). Consistent with this, most prefrontal interneurons have similar response properties to nearby pyramidal neurons (Rao et al., 1999), indicating that they form locally confined functional columns with balanced excitation and inhibition (Marino et al., 2005). All of these results suggest that, from early to late stages of visual processing, signals from opposite visual fields are processed relatively independently in different functional modules, compared with those from the same visual field that undergo highly competitive interaction through lateral inhibitions. Thus, the different neuronal modulation for targets presented within and across hemifields observed here is likely to inherit the division of contralateral and ipsilateral information in the upstream structures and be maintained through the local circuit organizations within the PFC.

Among many brain regions along the visual processing pathways, the PFC is most likely to limit the tracking capacity in the MOT task. We have previously demonstrated that prefrontal neurons consistently discriminate the visually identical objects during covert tracking, that the attenuation of the signals results in erroneous choice (Matsushima and Tanaka, 2012), and that the perturbation of the signals can alter the target selection (Matsushima and Tanaka, 2014a). It is also consistent with the widely accepted view that the PFC is crucial for performing tasks with high cognitive demands (Kane and Engle, 2002; Marois and Ivanoff, 2005), especially such as precise object recognition and discrimination (Duncan et al., 1995; Buschman and Miller, 2007), and with the fact that the bilateral advantage is observed only when irrelevant objects are presented along with the targets (Holt and Delvenne, 2014).

One might argue that the difference in neuronal response to the bilateral versus unilateral targets reported in this study was modest relative to that reported previously (Kadohisa et al., 2013). Although the large difference in experimental conditions might be the major cause, there might be several other possible reasons. First, the hemispheric independence might be largest in the PFC during the initial competition between the visual stimuli. Kadohisa et al. (2013) found that, when the target and distractor were presented alone in each hemifield, prefrontal activity was
initially biased toward the contralateral stimulus regardless of its identity and then evolved to predominantly encode the target regardless of its location. During the motion periods, however, monkeys had already identified and directed attention to the targets, while ignoring the other, task-irrelevant distractors. Second, the interference on the target processing from another target, addressed in the present study, might be stronger than that from the distractor examined previously (Kadohisa et al., 2013). Because distractors are known to impede the target processing further as they serve as a target in a training history (Bichot et al., 1996; Ipata et al., 2006; Kadohisa et al., 2013), the processing conflict might be maximized between the two, equally task-relevant targets to corrupt the independence of each hemifield. Taking all these factors into consideration, the present data appeared to be compatible with the previous observations, although our data provided a first evidence that the attentional enhancement of response to multiple targets depended on relative target locations, even after the priority over the distractors was fully established.

Parallel versus serial control of attention for multiple objects
Since Pylyshyn and Storm (1988) first devised the MOT task, several models of neuronal processing have been proposed. Although each model has unique properties, one of the major differences resides in whether attention is allocated to individual targets (Pylyshyn, 1989; Kahneman et al., 1992) or to a group of targets (Yantis, 1992). In support of the latter hypothesis, subjects sometimes have introspections that the multiple targets constitute a virtual polygon (Alvarez and Cavanagh, 2005). In the current study, however, we found that the same population of neurons exhibited enhanced activity and represented target locations in both the MOT and SOT trials, suggesting that attention might be separately directed toward each of the multiple targets. These results might be attributed to the lack of similarity or symmetry of target motions in our experimental conditions, under which perceptual grouping would be difficult (Koffka, 1935).

How is attention allocated to multiple objects? There is a longstanding debate between two hypotheses (Townsend, 1990). Attention might be focused on one location at a given moment and shifted sequentially (Treisman and Gelade, 1980; Wolfe, 1994), or alternatively, divided into multiple foci simultaneously for parallel processing (Duncan and Humphreys, 1989; Godijn and Theeuwes, 2003; Niebergall et al., 2011b). Although the comparable response variability between the MOT and SOT trials in the sampled neuronal population favored the latter hypothesis, the inconsistency across individual neurons prevented us from drawing any firm conclusion. Low firing rates of PFC neurons further complicated the analysis of oscillatory transition on the scale of a few tens or hundreds of milliseconds (Horowitz et al., 2009; Chakravarthi and VanRullen, 2011). Future studies would need more intricate investigations to address this issue.

References