Neuronal and Perceptual Effects of Selective Attention in the Primate Visual System

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August Universität zu Göttingen

vorgelegt von

Dipl.-Biol.
Robert Niebergall
aus Berlin

Göttingen 2009

D7

Referent: Prof. Dr. Stefan Treue

Korreferent: Prof. Dr. Julia Fischer

Tag der mündlichen Prüfung: 19.01.2010

Acknowledgments

The work included in this thesis would never have been possible without the tremendous support of my supervisors Stefan Treue and Julio Martinez-Trujillo. I am deeply grateful for their efforts to make this project happen, despite numerous difficulties. It was, and still is, a truly remarkable experience. I can hardly express my gratitude to Julio Martinez-Trujillo, who makes his lab a wonderful and inspirational place to work at. His creativity and expertise in research, his patience, and his motivational skills are exceptional. Further, I would like to thank Stefan Treue for the many opportunities he gave me during my time as his student and the great intellectual and financial support throughout the years. I also appreciate the kind support of Julia Fischer to evaluate this thesis.

I was fortunate to work with great colleagues, whose opinions and insights during discussions have been very influential. I owe great thanks to Paul Khayat for his help regarding the data recordings and analysis and, importantly, for Monday soccer. Florian Pieper has been tremendously helpful in many ways and has always shown incredible patience, even in moments where I could no longer hide my ignorance. I thank Therese Lennert for providing valuable comments on this manuscript, for our smooth cooperation regarding the animal care and setups, and for her pleasant character. I am grateful to Luigi and Sergio for cooperating. Walter Kucharski provided excellent technical assistance, and I thank Christine Pamplin for guiding me through the dungeons of McGill administration. I would also like to thank Tzvetomir Tzvetanov for his ‘mentorship’ during the first year of my doctoral studies, Ralf Brockhausen for solving computer problems and sharing the office, Steven Nuara and Dirk Prüsse for help regarding animal care issues, and Steffen Katzner for great teamwork as soccer coaches.

I thank my parents, Ursula and Gerhard, and my sister Anna-Lea for their unconditional support.

Last, but most important, I thank my wife Stephanie for sharing lives.

Contents

Chapter 1

Introduction

The central question of this thesis is how cognitive processes influence sensory visual information processing in the primate brain. The core of this work is composed of five manuscripts and one published article, each one representing a separate chapter. All manuscripts have been submitted and are currently at the first or second stage of the peer review process.

Four of these studies investigated how different aspects of selective attention modulate the neural mechanisms underlying visual information processing and perception in the primate brain. In order to conduct these studies, I recorded the activity of single cells and local field potentials in the visual cortex of awake, behaving macaque monkeys. The macaque monkey is a widely used animal model in visual and cognitive neuroscience, mainly because of the striking similarity of its visual system to that of humans, as well as its ability to learn and perform complex behavioral tasks.

Complementing the electrophysiological studies in the monkey, I conducted behavioral experiments in order to investigate how selective attention shapes the visual perception and response behavior of human subjects. The results of these experiments are reported in one article and one manuscript, both of which are included here.

The present chapter consists of three parts. In the first one, I will briefly outline the different processing stages of visual information, from the retina through the visual areas of the primate brain. The second part focuses on structural and functional properties of the motion sensitive area MT, the brain area targeted in all our electrophysiological experiments. The last part introduces the concept of selective attention, the central mechanism of the brain for the selection and modulation of behaviorally relevant sensory signals.

The main part of this work contains the original research manuscripts. Each one is preceded by a brief introduction, specifying the objective of the study.

1.1 The primate visual system

1.1.1 Signal processing in the retina

Light reflected by objects in the three-dimensional world is focused by the cornea and lens forming two-dimensional images on the retinal surface. A dense array of photoreceptors along the back surface of the retina converts such images into electrochemical signals (Tovee, 2008). Due to the spatial arrangement of the photoreceptors and their spectral and temporal properties the intensity of light can be encoded as a function of position (in two dimensions), wavelength and time. These primary signals are processed within a complex network of interneurons (horizontal, bipolar, and amacrine cells) located in several intermediate layers of the retina before they converge onto a population of approximately 1.5 million retinal ganglion cells of distinct morphological and functional features. Each ganglion cell responds to the onset of a light stimulus within a spatially restricted area of the retina, called the receptive field (RF). Stimulation of a location outside the RF does not modulate the cells' activity.

The concept of RFs is fundamental for understanding the functionality of sensory visual neurons and information processing in general. Ganglion cells have RFs with a characteristic center-surround organization (Schiller *et al.,* 1986). About half of the ganglion cell population is hyperpolarized (more negative membrane potential relative to resting state) by a light onset in their RF center ("off" ganglion cells), causing a suppression of the cells' spontaneous activity. The other half is depolarized (more positive membrane potential), resulting in a discharge rate higher than spontaneous activity ("on" ganglion cells). Changes in the activity of 'on' and 'off' center cells can signal local differences in light intensity, an essential feature for encoding luminance contrast. Other basic stimulus properties, such as spatial and temporal frequency, and color contrast are also encoded by circuits of retinal ganglion cells. Thus, visual information processing occurs as early as within the retina.

Besides functional RF properties, a second classification scheme for retinal ganglion cells is based on their morphology, their projection sites, and the type of information they encode (Shapley and Perry, 1986). Two major ganglion cell classes have been identified, which constitute the origin of two separate visual pathways. The vast majority (about 70-80%) belongs to the group of parvocellular (P) cells. These

cells are characterized by small RFs, low contrast sensitivity, slow axonal conduction velocities, and they are sensitive to high spatial and low temporal frequencies. They give sustained responses to a maintained stimulus and signal chromatic information, since they are selective for certain electromagnetic wavelengths. In contrast, ganglion cells of the magnocellular (M) pathway (about 10% of the total ganglion cell population), generally have large RFs, high contrast sensitivity, fast axonal conduction velocities, and are sensitive to high temporal and low spatial frequencies. They respond only transiently to a maintained stimulus and show poor selectivity for different wavelengths, rendering their signals achromatic. In short, M cells signal the presence of moving or changing patterns over a wide range of spatial scales and temporal frequencies, but leave the details to be carried by the P cells. These different response properties of retinal ganglion cells determine the functional properties of their projection neurons along the visual pathways.

1.1.2 Subcortical visual pathways

The axons of all retinal ganglion cells merge to form the bilateral optic nerves and leave the eye through the optic disc. The optic nerves from both eyes converge and cross at the optic chiasm. Here, axons originating from ganglion cells of the nasal (inner) part of each retina crossover to the contralateral hemisphere, while those from ganglion cells of the temporal (outer) part of the retina remain on the same side. The axons of ganglion cells representing the contralateral part of the visual field from one eye join the axons of ganglion cells representing the ipsilateral part of the visual field from the other eye. As a consequence, each hemisphere of the brain receives information only from the contralateral side of the visual field. The reorganized axon fibers form the optic tracts, which transmit the visual information of each visual hemifield to subcortical brain structures.

The lateral geniculate nucleus (LGN) in the thalamus is the major relay nucleus for inputs into visual cortex. About 90% of all retinal ganglion cell axons terminate here. The remaining 10% of ganglion cell fibres project to other thalamic structures, including a pathway from the retina through the superior colliculus (SC) to the pulvinar, which in turn has reciprocal connections with several extrastriate cortical areas (Cowey and Stoerig, 1991). The LGN is composed of six main layers, which maintain the separation of the inputs from the P and M pathways. The upper four layers receive inputs from P retinal ganglion cells, while the lower two receive inputs

from M retinal ganglion cells. A third population of very small (koniocellular) neurons is present in the LGN, constituting a third independent processing stream. This koniocellular (K) stream receives inputs from the SC as well as from axons originating in the retina, which convey a chromatic (blue-on, yellow-off colour-opponent) signal (Hendry and Yoshioka, 1994). However, neither the locations of these LGN cells nor the type of retinal ganglion cell providing their input has yet been determined with certainty. In the layered structure of the LGN, the inputs from the two eyes are aligned but remain physically segregated. Consequently, LGN neurons are monocular. They have small RFs with a concentric centre–surround structure. Individual LGN cells relay the outputs of retinal ganglion cells in approximately 1:1 fashion to the primary visual cortex.

1.1.3 Primary visual cortex

The primary visual cortex (V1) is located in the posterior part of the occipital cortex. Due to its layered or striped appearance in cross sections stained for cytochrome-c-oxidase it is also referred to as 'striate' cortex (Horton and Hubel, 1981). Neurons in the M and P layers of the LGN project to different layers in V1, producing a precise topographic representation of the visual field on the cortical surface. Great emphasis is given on central (foveal) vision, indicated by the fact that about half of the cortical space in V1 is devoted to the central 5° of the visual field (a phenomenon known as cortical magnification). V1 is composed of six layers, which differ in cell density, connectivity, and many other characteristics. Each layer contains alternating ocular dominance stripes receiving inputs exclusively from only one eye, indicating that inputs from both eyes remain segregated at this stage. The inputs from the M and P pathway also terminate in different layers of V1. Even though the segregation of these two pathways, which was evident at the retinal and subcortical level, is maintained (Livingstone and Hubel, 1988), a significant interchange of the two pathways in the first few synapses in V1 has recently been demonstrated (Sincich *et al.,* 2004).

In order to extract the visual information transmitted through the retino-subcortical pathways the number of neurons available to process these inputs is dramatically expanded, resulting in hundreds of cortical neurons for each LGN input (Schein and de Monasterio, 1987). V1 neurons also show many new tuning properties relative to those of LGN neurons. They are selective for orientation, motion direction, color, and binocular disparity (Livingstone and Hubel, 1988). They also show

sensitivity to contrast gradients produced by bar stimuli, edges, and borders and are therefore well suited to encode simple stimulus features. Their RFs are small, providing detailed feature information at a high spatial resolution (Barlow *et al.,* 1967; Cumming, 2002). In summary, V1 extracts many features of the visual scene at a fine scale before selective aspects of this information are routed to different areas in extrastriate cortex, which subserve more specialized analyses.

1.1.4 Visual information processing in extrastriate cortical areas

The outputs from V1 are composed of at least two anatomically and functionally segregated processing streams, the *ventral* and the *dorsal* pathway (Ungerleider and Mishkin, 1982). Areas constituting the *ventral* pathway (V1, V2, V4, TEO, IT) are involved in the identification of shapes and color and are mostly found in the posterior/inferior (occipito-temporal) part of the brain. According to its functional specialization this pathway is commonly called the "what" pathway. Areas forming the *dorsal* pathway (V1, V2, V3, MT, MST, FST, STP, VIP, LIP, 7A) are involved in the localization of objects as well as the perception of motion. They are mostly found in the posterior/superior (occipito-parietal) part of the brain. This pathway is commonly referred to as the "where" pathway.

Even though the two streams operate relatively independently, there is ‘cross-talk’ between them at several levels (DeYoe and Van Essen, 1988). The classical view of the *dorsal* and *ventral* streams mediating different visual attributes (Ungerleider and Mishkin, 1982) has more recently been challenged by the proposal that both streams might make use of similar sets of visual attributes, but for different behavioral goals (Goodale and Milner, 1992; Rizzolatti and Matelli, 2003).

Most extrastriate visual areas receive almost all of their inputs, either directly or indirectly, from V1. In addition, most connections between visual areas consist of both feedforward and feedback projections, indicating that there is a high degree of interactive processing. As one ascends the visual hierarchy, neurons have progressively larger RFs and respond to stimuli of greater complexity. Early visual areas preferentially encode simple stimulus features, such as oriented lines (V1), whereas higher visual cortical areas preferentially respond to increasingly complex stimulus attributes. For instance, neurons in inferotemporal cortex (IT) are selective for different shapes of objects (Pasupathy, 2006), or even individual faces (Leopold *et al.,* 2006), revealing a high degree of holistic information processing.

1.2 Middle temporal visual area (MT)

During all electrophysiological experiments reported here, we exclusively targeted the middle temporal area (MT), also known as V5. MT is common to all primates (Kaas and Lyon, 2001), which is one of many reasons why MT has been the area of choice in numerous studies investigating the influences of cognitive processes on sensory information processing in single neurons. In the following section the properties of area MT most relevant to the present studies will be reviewed.

1.2.1 Location and structure

MT is part of the *dorsal* pathway and is located in the posterior bank of the superior temporal sulcus (STS). Like V1, MT is retinotopically organized (Van Essen *et al.,* 1981), with each hemisphere containing a complete map of the contralateral visual hemifield and a small portion of the ipsilateral field close to the vertical meridian (Desimone and Ungerleider, 1986). Foveal vision is markedly emphasized, with the central 15˚ of the visual field occupying over half of MT's surface area (Van Essen *et al*., 1981). There is also a biased representation of the lower visual field quadrant. Foveal vision and the lower visual field are represented in the lateral part of MT, while larger eccentricities and the upper visual field are represented more medially (Maunsell and Van Essen, 1987). The vast majority of neurons in area MT show strong selectivity for processing of motion direction and speed (see also '*Response properties of MT neurons',* pp. 8). Neurons sharing similar direction preferences are clustered in columns oriented perpendicular to the cortical surface (Albright *et al.,* 1984). All directions are uniformly represented in MT neurons with motion direction preference changing gradually in adjacent columns.

1.2.2 Connectivity

MT represents an intermediate stage within the hierarchy of visual information processing. It receives feedforward inputs from multiple cortical areas, including V1, V2, V3, V3A, VP, and PIP (Felleman and Van Essen, 1991; Maunsell and van Essen, 1983). The main input to MT, however, is a mainly magnocellular (M) projection, originating from direction and speed selective complex cells in V1 (Movshon and Newsome, 1996). Even though the cortical inputs to MT predominate, some MT neurons remain both visually responsive and even direction-selective after removal or

inactivation of V1. This residual functionality might derive from callosal connections from the intact hemisphere (Girard *et al.,* 1992), or direct subcortical inputs from the SC (Rodman *et al.,* 1990) and koniocellular neurons of the LGN (Nassi and Callaway, 2006; Sincich *et al.*, 2004). Nevertheless, MT and the *dorsal* stream rely heavily on visual information provided by the magnocellular pathway. This is demonstrated by the fact that reversible inactivation of the M layers of the LGN nearly completely abolishes the visual responsiveness of MT neurons, whereas P-layer inactivation has a much smaller, though measurable, effect (Maunsell *et al.,* 1990).

The main cortical target regions for feedforward projections arising in MT are its neighboring areas FST and MST in the STS, parietal lobe areas such as VIP, LIP and 7a and also frontal lobe areas such as FEF and the dorsolateral prefrontal cortex. In addition, MT also has multiple feedback projections to cortical (V1, V2, V3A) and numerous subcortical (e.g. dorsal LGN, pulvinar, and SC) regions. For example, MT inactivation affects orientation and direction selectivity in V2 neurons, indicating that MT feedback projections influence these neurons' RF properties (Gattass *et al.*, 2005).

1.2.3 Receptive field structure of MT neurons

The classical RF diameters of MT neurons are about equal to their eccentricity and therefore about 10 times larger than the diameter of their V1 inputs (Born and Bradley, 2005). As a consequence, considerable spatial pooling arises in the formation of MT cells' RFs. About half of the neurons in MT have RFs with direction selective antagonistic surrounds, which on average spread across an area about three times the size of the classical RF diameter (Allman *et al.,* 1985; Raiguel *et al.,* 1995; Tanaka *et al.*, 1986). A stimulus extending outside the classical RF of an MT neuron with antagonistic surround suppresses the neurons response. This effect is strongest, when the stimulus motion in the surround represents the neurons preferred direction. Surround suppression is contrast dependent and vanishes at low contrast (Pack *et al.,* 2005). MT neurons will therefore respond better to a large stimulus with low contrast than to one with high contrast. In the macaque, neurons with antagonistic surround RFs are more common in the output layers, whereas those lacking antagonistic surrounds are found predominantly in the input layers (Raiguel *et al.*, 1995).

1.2.4 Response properties of MT neurons

Like other areas of the superior temporal sulcus (MST, FST) MT contains many cells sensitive to the direction of motion and the speed of a stimulus (Dubner and Zeki, 1971; Maunsell and Van Essen, 1983a). Mapping the responses of MT cells with stimuli of different motion directions inside the RF typically reveals a Gaussian-shaped tuning curve. The peak of the curve, representing the strongest response, is centered on the neurons preferred direction. In contrast, motion directions opposite to the preferred direction usually produce weaker responses. This direction is commonly referred to as 'anti-preferred'- or 'null'-direction. A measure for selectivity in the directional tuning is the bandwidth, which is defined as the width of the tuning curve at half of the difference between the response to the preferred and antipreferred directions (Maunsell and Van Essen, 1983a). On average, response increments for motion in the neuron's preferred direction are about four times the magnitude of response decrements for motion in its null direction (Snowden *et al.,* 1991).

Most MT cells are also tuned to motion speed (Maunsell and Van Essen, 1983a). They are typically bandpass-tuned with a preference for intermediate speeds, while slower or faster speeds lead to response decreases. In addition, MT is sensitive to other aspects of visual information, such as stimulus orientation (Albright *et al*., 1984), binocular disparity (Maunsell and Van Essen, 1983b) and the direction of smooth pursuit eye movements (Komatsu and Wurtz, 1988).

The tuning properties of MT neurons suggest that this area plays a key role in the perception of visual motion signals. Indeed, several studies have demonstrated a close link between MT neuronal activity (or its absence) and perceptual experience. For example, macaque monkeys with lesions in MT show impaired direction discrimination performance, while their contrast perception remains unaffected (Newsome and Pare, 1988; Pasternak and Merigan, 1994). Furthermore, the monkeys' decision during direction discrimination can be biased by electrical stimulation of an MT direction column (Salzman *et al.,* 1992), and there is a high degree of correlation between the neural threshold and the behavioral threshold for direction discrimination (Britten *et al.,* 1992). This close relationship between neuronal and psychophysical threshold persists even when both vary across a recording session (Zohary *et al.,* 1994).

1.3. Mechanisms of attention

Vision is far more than a passive reflection of our environment. It is a highly active process by which visual information on its way from the retinas through visual cortex is extracted and continuously modified. The specialized circuits of retinal ganglion cells, extracting basic sensory cues even before the visual signals leave the eye, are only one of many striking examples of the visual systems' processing strategies. Even though our visual experience appears to be a coherent representation of the visual world, only a small amount of the visual information entering the eyes is consciously perceived and stored in memory (Lamme, 2003). This indicates that the processing capacity of the visual system is limited and that multiple aspects of the visual information must compete for access to this system (Desimone and Duncan, 1995).

A mechanism for the selection and modulation of behaviorally relevant sensory information in the brain is attention. It acts as a filter, selecting those parts of the information that will subsequently be available for and amplified by higher-level processing in different areas along the visual pathways. Throughout this work, the term attention is defined as the selective filtering and modulation of sensory information according to its behavioral relevance. Since it is the central topic of this thesis, those aspects of attention relevant to the studies included here will be briefly summarized in the following sections.

1.3.1 Perceptual correlates of attention

A number of classical behavioral studies have compared visual attention to a moving 'spotlight', which can be allocated to different positions within the visual field (Eriksen and St James, 1986; Posner, 1980). Information within the attentional spotlight is selectively highlighted, resulting in enhanced sensitivity (e.g., Carrasco *et al.,* 2004), faster detection (e.g., Eriksen and St James, 1986) and more precise identification (e.g., Henderson and Macquistan, 1993) in response to attended items.

One property of the attentional 'spotlight' is that it can be covertly oriented (covert attention), i.e., it is freely directed upon any selected location in the visual field, unaccompanied by eye movements. In a typical attention task, subjects maintain fixation on one point in the visual field while attending to eccentric stimuli, separating the mechanisms related to eye movements from those related to attention. However, under natural conditions, covert shifts of attention are often followed by

matching eye movements. Due to this latter relationship, some studies have suggested a close relationship between brain systems controlling eye movements and those controlling shifts of attention (Moore and Armstrong, 2003; Rizzolatti *et al.,* 1987).

Within the domain of spatial attention two sub categories have been distinguished (Ling and Carrasco, 2006). Involuntary attention (also referred to as 'exogenous', 'bottom-up' or 'stimulus-driven') has fast build-up rates (about 100 ms) and is driven by virtue of the relative salience of the target. Enhanced salience might arise from the sudden appearance of the object in the scene or by a particular feature that makes it 'pop-out' from an array of other stimuli (Desimone and Duncan, 1995). In contrast, voluntary attention (also referred to as 'endogenous', 'top-down' or 'goal-directed') is typically characterized by slower time courses (about 300 ms) and the willful selection of relevant objects by the subject.

Although *spatial attention* is the most widely studied attentional mechanism, attentional selection is not restricted to the spatial location of objects. For instance, attending to a particular stimulus feature, e.g. a particular color or motion direction enhances processing of the attended feature across the entire visual scene, independent of the spatial focus of attention. This mechanism is known as *feature-based attention* (Saenz *et al.,* 2002; Treue and Martinez Trujillo, 1999; Zhang and Luck, 2009). Since the perceptual enhancement of features is independent from spatial locations, feature-based attention is particularly useful in visual search tasks where a predefined target item embedded in a display of distractors needs to be selected (Bichot *et al.,* 2005). A third attentional mechanism, *object-based attention*, refers to the enhanced processing of all features of an object, even if they are irrelevant and attention is directed to only one of them (Blaser *et al.,* 2000; Egly *et al.,* 1994; O'Craven *et al.,* 1999; Roelfsema *et al.,* 1998).

1.3.2 Neural correlates of attention

In the last decades, neuroscientists have tried to disentangle the basis of the modulatory effects of attention at the level of single neurons in different brain areas, using electrophysiological methods in awake, behaving macaque monkeys. The activity of neurons in the cortex of these animals can be monitored via extracellular single-cell- and local field potential (LFP) recordings, while the monkeys perform visual tasks that require the allocation of attention to different spatial locations or stimulus features. Many of these experiments have shown that directing attention to a

stimulus located inside a neurons receptive field modulates the neurons' response (Moran and Desimone, 1985). If the features of the attended stimulus match the neurons preference, the neuron typically responds by increasing its spiking activity (firing rate), whereas an attended stimulus with non-preferred features typically reduces (and sometimes entirely shuts off) the spiking activity.

The strongly selective nature of attention becomes evident when two stimuli with opposing features are presented in the RF. For instance, two stimuli moving in different directions presented inside an MT neuron's RF produce a response that resembles a weighted average of the responses to the individual stimuli when attention is directed outside the RF (Treue and Maunsell, 1996). However, switching attention between the two stimuli inside the RF causes a suppression of the neuron's response when attending to the non-preferred stimulus and a response enhancement when the preferred stimulus is attended. This indicates that attention specifically enhances the contribution of the attended stimulus to the response at the expense of unattended stimuli.

In contrast, the effects of feature-based attention do not seem to depend on the spatial location of the attentional focus. For example, if an animal attends to a moving stimulus located outside a neurons RF, neuronal responses to a second, spatially unattended stimulus inside the RF can be enhanced if that stimulus moves in the same direction as the attended stimulus and if the neuron preferentially encodes the attended stimulus direction (Treue and Martinez Trujillo, 1999). Under such circumstances, attention to a particular stimulus feature, such as the direction of motion, enhances the response of cells sharing feature preferences similar to those of the attended stimulus, independent of the location of spatial attention relative to the neurons' RFs.

The modulatory influence of attention on neuronal responses increases as one moves up the hierarchy of visual information processing (Maunsell and Treue, 2006). Nevertheless, there is evidence for attentional modulation of neuronal responses in subcortical structures such as the LGN (e.g., McAlonan *et al.,* 2008), the adjacent thalamic reticular nucleus (TRN) (McAlonan *et al.*, 2008), and the SC (e.g., Fecteau and Munoz, 2005). The effects of attention have also been reported in striate cortex (e.g., Motter, 1993), a large number of extrastriate cortical areas in both *ventral* stream areas (e.g., Spitzer *et al.*, 1988) and *dorsal* stream areas (e.g., Treue and Martinez Trujillo, 1999) as well as in areas of the frontal lobe (e.g., Buschman and

Miller, 2007). In humans, functional magnetic resonance imaging (fMRI) studies have also reported physiological correlates of the effects of attention (Silver and Kastner, 2009). Cortical areas with retinotopic representations of visual space (Gardner *et al.*, 2008) show focal enhancement of BOLD signals that correlate with covert shifts of the attentional spotlight toward the corresponding regions of the visual field (Brefczynski and DeYoe, 1999). Similar findings have been reported in studies of brain electrical activity using Event Related Potentials (ERPs) (Hillyard and Anllo-Vento, 1998). In general, these studies demonstrate that the processing of visual signals in the human brain is strongly modulated by attention.

1.3.3 Models of attentional signal enhancement

Although there is agreement that attention changes the neural response to unchanged sensory conditions, the underlying mechanisms of this attentional modulation are still under debate. Two main models exist: the 'response gain model' and the 'contrast gain model'. Both models describe an increase in the responses of neurons encoding the attended stimulus feature or locations. However, the models differ in the mechanism by which this enhancement is achieved. The 'response gain model' ascribes the effects of attention to a 'scaling' of neuronal responses by a uniform factor (multiplicative modulation), modulating the sensitivity of a neuron to any input signal (McAdams and Maunsell, 1999). This operation can be compared to the effect of an amplifier that increases the gain of incoming signals. The 'contrast gain model', on the other hand, proposes that attention increases the relative salience of a stimulus in a manner similar to changes in stimulus contrast. For example, the magnitude of the attentional enhancement changes as a function of the luminance contrast of a moving stimulus relative to a dark background (Martinez-Trujillo and Treue, 2002; see also Reynolds *et al.*, 2000 for similar results in V4). The difference to the previous model is that here, the signal amplification by attention depends on the stimulus contrast. For very high or very low contrasts, the signal increase by attention is smaller than for intermediate contrasts.

To better understand the difference between the two models one could consider the effects of attention on the contrast response function of a neuron to a stimulus. In visual neurons, it has been widely documented that the shape of this function is sigmoidal: the neuron increases its firing rate as a function of increasing stimulus contrast until it reaches a saturation point where further increases in contrast do not

produce higher firing rates. A multiplicative modulation would predict that when a stimulus is attended, the entire function is multiplied by a certain factor (contrast independent modulation). On the other hand, the contrast gain model predicts that when the stimulus is attended the increase is only evident for intermediate contrast values but not for the lowest and highest contrasts. The latter would lead to a shift of the sigmoidal contrast–response function along the contrast axis.

The implication of the contrast gain model is that the mechanisms of attention may be intermingled with (and perhaps indistinguishable from) the mechanisms that determine the saliency/contrast of visual stimuli. Although several models have been proposed over the last year (Ghose, 2009; Lee & Maunsell, 2009; Reynolds & Heeger, 2009), no conclusive data have yet been reported that could decisively distinguish between the models and solve this debate. The studies in chapter 2.3 and chapter 2.4 address this issue, clarifying our understanding of the mechanisms of attentional signal enhancement.

1.3.4 Attentional selection of multiple objects

Classical theories of attention typically assumed a single attentional focus for the selection of relevant information (Eriksen and St James, 1986; Posner, 1980). However, it is evident that many everyday activities, such as team sports, video games, or navigating through busy traffic, require attention to multiple regions of interest. More recent studies indeed showed that the spatial focus of attention can be split, such that attention is simultaneously directed to multiple locations in the visual field, excluding interspersed regions (Awh and Pashler, 2000; McMains and Somers, 2004; Morawetz *et al.*, 2007; Muller *et al.*, 2003). Using fMRI, McMains and Somers (2004) were able to image two separate peaks of activation in striate and extrastriate cortices corresponding to two separate attended locations. However, the issue of dividing attentional resources remains controversial since contradicting results have also been reported, showing that the influence of distractor stimuli in between the attended locations might not be excluded (Heinze *et al.*, 1994; Muller *et al.*, 2003a).

The central question of whether and how attention can be divided has also been investigated in tasks that closely resemble real life situations, such as the active tracking of multiple moving objects. In those tasks, subjects typically track multiple predefined target stimuli embedded within a set of distractors, randomly moving across a computer screen. It has been demonstrated that human observers can track up

to five targets simultaneously (Pylyshyn and Storm, 1988). While it is generally accepted that the ability to keep track of multiple moving objects critically relies on attention (Allen *et al.*, 2004; Yantis, 1992), the mechanisms by which attention connects to the different targets and maintains these connections as the targets change location have yet to be revealed. The most prominent model, multifocal attention, assumes that each target attracts an independent focus of attention, following the targets as they move. This strategy relies only on basic properties of attention, i.e., the selection of individual stimuli according to their relevance, but requires that attention can deploy more than one focus.

If this model is correct, the question arises how the individual attentional foci enhance the representation of the tracked targets. One line of evidence suggests that attention does not enhance the tracked objects themselves, but rather suppresses the influence of the distractors (Pylyshyn, 2006; Pylyshyn *et al.*, 2008). A recent event-related potential (ERP) study (Drew *et al.*, 2009), however, argues in favor of an attentional mechanism that enhances the representation of the tracking targets without any evidence for distractor suppression below the level of the background. Although several studies have addressed different aspects of attentional selection during multiple-object tracking, using fMRI (Culham *et al.*, 1998; Howe *et al.*, 2009), ERP (Drew *et al*., 2009), and psychophysical methods (for review see Cavanagh and Alvarez, 2005) substantial questions regarding the precise mechanisms of target enhancement and distractor exclusion as well as the question of a single versus multiple attentional spotlights remain largely unresolved. To clarify these issues we recorded the activity of single neurons in area MT of the macaque during attentional tracking tasks. The results of this study are presented in chapter 2.1.

Original articles and manuscripts

This chapter contains the following articles and manuscripts:

- Niebergall, R., Khayat, P.S., Treue, S., Martinez-Trujillo, J.C. Splitting and zooming the focus of attention in primate visual cortex during multiple-object tracking. Prepared for submission. Author contributions: RN conceived and conducted the experiments, analyzed the data and wrote the manuscript.
- Niebergall, R., Huang, L., Martinez-Trujillo, J.C. Similar perceptual costs for dividing attention between retina- and space-centered targets in humans. Submitted. Author contributions: RN conceived and conducted the experiments, analyzed the data and wrote the manuscript.
- Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C., (2010). Attention Differentially Modulates Similar Neuronal Responses Evoked by Varying Contrast and Direction Stimuli in Area MT. *The Journal of Neuroscience*, 30(6): 2188–2197. Author contributions: RN conducted the experiments.
- Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C., (2010). Frequency-Dependent Attentional Modulation of Local Field Potential Signals in Macaque Area MT. *The Journal of Neuroscience*, 30(20): 7037–7048. Author contributions: RN conducted the experiments.
- Tzvetanov, T., Womelsdorf, T., Niebergall. R., Treue, S., (2006). Feature-based attention influences contextual interactions during motion repulsion. *Vision Research*, 46(21): 3651–3658. Author contributions: RN conducted the experiments and analyzed the data.
- Sachs, A.J., Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C. Contribution of spike timing to contrast and motion direction coding by single neurons in macaque area MT. Prepared for submission. Author contributions: RN conducted the experiments.

2.1 Splitting and zooming the focus of attention in primate visual cortex during multiple-object tracking

Classical psychophysical studies have described spatial attention as a spotlight or unitary focus, where processing resources are concentrated. Over the past years, some studies have provided evidence that this focus can 'expand', or 'contract', depending on the size of the attended region and demands of the task. In addition, more recent evidence suggests that attention can be divided into multiple foci, each one independently enhancing the processing of information from different regions of the visual field. Previous studies addressing this issue have used functional imaging (fMRI) and event related potentials (ERP) to show enhanced activity within brain maps representing spatially separated regions of the visual field during divided attention tasks. However, these brain mapping methods register the activity of millions of neurons at a given location, lacking the spatial resolution necessary to test whether attentional resources can be divided at the level of a single neurons' receptive field.

In this work we examined this and other related questions. We trained two rhesus monkeys to simultaneously direct attention to two spatially separated moving random dot patterns (RDPs) while ignoring a third RDP. The latter was positioned at the center of the recorded MT neurons receptive field in between the two attended RDPs. When the separation between the attended RDPs was larger than the size of the neurons' receptive field the response to the ignored stimulus remained unchanged. This result suggests a split of the attentional spotlight into two separate foci that selectively modulated responses to the peripheral stimuli. However, positioning the attended stimulus pair inside the receptive field caused an enhancement in response to both the attended stimuli in the periphery and the ignored interspersed stimulus. Thus, attentional resources were spread across all stimuli inside the receptive field, indicative of a single spotlight of attention that increased in size to match the receptive field size.

Overall, these results show that the distance between attended stimuli relative to the receptive field size of neurons within areas such as MT determines whether attentional modulation of single cell responses is compatible with the existence of a single or multiple attentional spotlights.

Splitting and zooming the focus of attention in primate visual cortex during multiple-object tracking

Robert Niebergall[1,2], Paul S. Khayat[1], Stefan Treue[2,3], and Julio C. Martinez-Trujillo[1]

[1]*Cognitive Neurophysiology Laboratory, Dept. of Physiology, McGill University, 3655 Promenade Sir William Osler, Montreal, QC, H3G 1Y6, Canada.*

[2]*Cognitive Neuroscience Laboratory, German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany.*

[3]*Bernstein Center for Computational Neuroscience, Bunsenstrasse 10, 37073 Goettingen, Germany.*

Abstract

Simultaneous tracking of multiple moving objects imposes a challenge to our visual system, highly specialized for foveal vision. We studied the brain mechanisms underlying multiple-object tracking by recording responses of neurons in visual area MT of monkeys during a tracking task. Neurons responded more strongly when the animals attentively tracked two moving objects than when ignoring them, demonstrating that attention enhances neural representations of tracked objects. This enhancement split into multiple foci, when the separation between tracked objects was larger than the size of the neurons' receptive field, or zoomed out over tracked objects and intermediate distracters when objects fell within the same receptive field. These results demonstrate a brain mechanism that adaptively distributes processing resources amongst visual neurons encoding objects' representations during tracking.

Multiple-object tracking (MOT) has become an intensive area of research in human and computer vision (*1-11)*. Common activities, such as conducting active surveillance at airports, or playing team sports rely on the ability to simultaneously track multiple objects that move in space and across our retinas in the presence of distracters (Fig. 1A). How does our visual system, apparently evolved to align the high-resolution fovea (*12*) with one object at a time, accomplish such a challenge? Making rapid saccades from one moving object to another is costly and inefficient, because the brain suppresses vision during saccades (*13*). Instead, it has been proposed that covert orienting of visuospatial attention underlies MOT (*1-10*). Attention could, independently from gaze direction, enhance representations of tracked objects across retinotopically-organized maps of visual space in cortical areas of the primate brain.

Visuospatial attention has been described as a *spotlight* that enhances the processing of a behaviorally relevant object (*14, 15*), as a *zoom lens* that adjusts its size to enhance the processing of several objects over a continuous region of space (*16-18*) and as a flexible resource that can split into *multiple foci*, selectively enhancing the processing of tracked objects and excluding interspersed distracters (*19-21*). Functional imaging (*1,2*) and event-related potential (*3*) studies in humans have suggested that during MOT attention modulates the activity of neurons in visual cortex. This hypothesis, however, remains untested. Furthermore, it is unknown how attentional resources are deployed during tracking.

We investigated these issues by recording single cell responses from the middle temporal area (MT) of two monkeys (Macaca mulatta) performing a MOT task. Macaques, like humans, rely on MOT during their interaction with the environment (e.g., an alpha male tracking his herd members). The similarity in structure and function of area MT amongst primates (*22*), the retinotopic organization of MT neurons RFs (*23*), and the fundamental role of MT neurons in the processing of visual motion make the area an ideal target for our recordings (*24*). Additionally, functional imaging studies have reported increased metabolic activity within the human MT/V5 complex during MOT (*1, 2*).

Throughout experimental trials, the animals sat in front of a screen where a fixation spot and three random dot patterns (RDPs) were presented. One RDP was positioned inside the recorded neuron's RF (RF-pattern), and the other two translated

across the screen following parallel trajectories alongside, but never entering the RF (MOT-targets). Dots in the three RDPs locally moved in the neuron's preferred direction and at the preferred speed (Fig. 1B) (*25*).

Neuronal responses and behavioral performance were recorded during three tasks. First, when animals attended to the fixation spot and detected a transient change in its luminance (*fixation*) (*24*). Second, when animals tracked the MOT-targets and detected a transient change in one of the patterns dots' speed (*MOT*). Third, when animals attended to the RF-pattern and detected a similar change (*RF-task*).

We examined whether during *MOT* the animals used a single spotlight of visuospatial attention that rapidly switched from one target to another (*10*). Performance data indicate that the animals would have had to switch attention at a rate of 16 Hz (60 ms per switch), or higher to reach the observed performance levels (Fig. S3C). The reported time for macaques to switch attention between two stimuli is about 150-180 ms (~6.6-5.5 Hz) (*26, 27*). In humans, it is about 150 ms (~6.6 Hz) (*4*). We therefore discard this possibility and conclude that the animals tracked the MOT-targets in parallel (*4, 5*).

We reasoned that parallel tracking could be accomplished in two different manners. Either the animals tracked the MOT-targets with a large spotlight of attention (*zoom lens*) (*16-18*), or with two individual spotlights (*multifocal attention*) (*19-21*). Tracking with a *zoom lens* predicts that when MOT-targets pass alongside the RF-pattern, attention enters the RF and enhances neuronal responses relative to *fixation* (Fig. 1C) (*24, 28*). On the other hand, tracking with multiple foci predicts that during *MOT* attention remains outside the RF and responses are similar to those during *fixation* (Fig. 1D).

We recorded responses from 108 MT neurons in two animals, estimated each neuron's RF diameter, divided it into five regions, and added two additional regions on each side (Fig. 1E) (*25*). This allowed us to determine each neuron's average response along the MOT-targets' trajectory. For the neuron in Fig. 1E, responses in the *RF-task* (green) where higher than during *fixation* (blue), demonstrating that directing attention into the neuron's RF enhances its response (*24, 28*). More importantly, responses during *MOT* (red) were similar to those during *fixation* (blue), and considerably lower than during the *RF-task* (green), suggesting that along the MOT-targets' trajectory attention never entered the RF area.

For the majority of recorded neurons (Fig. 1F), and in all tasks, responses were similar across regions ($P = 0.603$, $P = 0.826$, and $P = 0.799$, for *fixation*, *MOT* and *RF-task*, one-way ANOVA with region as main factor). Importantly, pooling responses across regions did not reveal differences between *MOT* and *fixation* ($P = 0.153$, paired t-test) but a strong response increase in the *RF-task* relative to both, *MOT* ($P < 0.0001$, paired t-test), and *fixation* ($P < 0.0001$, paired t-test). Since we did not observe any response suppression during *MOT* or *fixation* when the MOT-targets passed nearby the RF (Fig. S6) (*29*), these results cannot be explained by the patterns entering the RF's inhibitory surround and counteracting a response increase produced by attention (*30*). Instead, our results show that attention split into two *foci* corresponding to the tracked targets and excluding the RF area (*multifocal attention*).

A distinctive feature of MOT is that targets constantly change retinal position, activating different neurons across retinotopically-organized maps in visual cortex. In order to enhance the responses of these neurons, attention must, together with the targets dynamically change position within the map. We tested this hypothesis by recording responses of 42 additional MT units in both animals during *MOT* and *fixation*, after removing the RF-pattern and decreasing the distance between the MOT-targets. This caused the MOT-targets to pass through the neurons' RF, producing a response increase ($P < 0.0001$, *MOT*, and $P = 0.0354$, *fixation*, one-way ANOVA, Fig. 2B).

The cell example in Fig. 2A and the population data in Fig. 2B show that when the MOT-targets were outside the RF (leftmost region) responses during *MOT* and *fixation* were similar ($P = 0.32$, paired t-test for the population data). However, when the stimuli were at the RF center, responses during *MOT* were on average 27% stronger than during *fixation* ($P = 0.00037$, paired t-test). This difference decreased again when the MOT-targets moved away from the RF, demonstrating that during *MOT* attention entered and abandoned the RF with the tracked targets.

Decreasing the distance between the MOT-targets produced a counterintuitive effect: a decrease in *MOT* performance relative to when MOT-targets were farther apart (Fig. 3). Interestingly, a similar effect has been reported in humans (*6*). Our results may have revealed its cause. We reasoned that when the separation between the MOT-targets was larger than the RF diameter (Fig. 1E), each target activated a unique and different population of neurons within the MT retinotopic map (*22*). This

allows independent coding of each MOT-target's features, and for multifocal attention to produce two independent foci of response enhancement within the map. On the other hand, when the distance between the targets is smaller than the RF diameter, they no longer activate two independent neuronal populations within the map, because some neurons have RFs including both targets. Here, attention may spread (like a *zoom lens*) to include both tracked targets and the region in between (*17*). This spread would explain the observed decrease in performance (*16, 17*).

In order to test the latter hypothesis we repeated our previous measurements after adding the RF-pattern (see Fig. 4A for a single cell example). The population data in Fig. 4B show that during *MOT*, responses at the RF center were significantly larger with (solid red line) than without (dashed red line) the RF-pattern ($P = 0.0049$, paired t-test), indicating that the RF-pattern was not excluded from the response but substantially contributed to it. Moreover, responses to the three stimuli in the RF were 27% stronger during *MOT* relative to *fixation* (blue), suggesting that attention enhanced responses to all stimuli ($P < 0.0001$, paired t-test).

Finally, if during *MOT* attention *zooms out* over the three stimuli, it will spread over a larger area relative to the *RF-task*, where attention was tightly focused on the RF-pattern. This spread will decrease the response in the former relative to the latter task (*17*). Indeed, population responses to the same three stimuli inside the RF were 13% smaller during *MOT* than during the *RF-task* ($P = 0.0068$, paired t-test). In general, these findings suggest that the attentional focus can vary in size, as predicted by the *zoom lens* hypothesis (*16, 17*).

In conclusion, during multiple-object tracking visuospatial attention enhances the representations of tracked objects within retinotopic maps in primate visual cortex. For targets separated by more than a RF diameter, this enhancement is achieved by a split of the attentional spotlight into multiple foci. For targets falling into the same RF, a single attentional spotlight spreads to cover all tracked stimuli and the area in between. This dichotomy could reconcile the apparent contradictions between previous studies providing evidence in favor of *multifocal* attention (*19-21*), and those supporting the *zoom lens* hypothesis (*16-18*). Our finding reveals a highly adaptive attentional system in the primate brain that flexibly distributes processing resources in space in order to match the complexity of the visual environment.

References and Notes

1. J. C. Culham *et al*., *J. Neurophysiol*. **80**, 2657 (1998).
2. P. D. Howe, T. S. Horowitz, I. Akos Morocz, J. Wolfe, M. S. Livingstone, *J. Vis*. **9**, 1 (2009).
3. T. Drew, A. W. McCollough, T. S. Horowitz, E. K. Vogel, *Psychon. Bull. Rev*. **16,** 411 (2009).
4. P. Cavanagh, G. A. Alvarez, *Trends Cogn. Sci*. **9**, 349 (2005).
5. Z. W. Pylyshyn, R. W. Storm, *Spat. Vis*. **3**, 179 (1988).
6. W. M. Shim, G. A. Alvarez, Y. V. Jiang, *Psychon. Bull. Rev*. **15**, 390 (2008).
7. G. A. Alvarez, P. Cavanagh, *Psychol. Sci*. **16**, 637 (2005).
8. C. R. Sears, Z. W. Pylyshyn, *Can. J. Exp. Psychol*. **54**, 1 (2000).
9. S. Yantis, *Cognit. Psychol*. **24**, 295 (1992).
10. L. Oksama, J. Hyona, *Cognit. Psychol*. **56**, 237 (2008).
11. S. W. Joo, R. A. Chellappa, *IEEE Trans. Image Process*. **16**, 2849 (2007).
12. A. E. Hendrickson, *Invest. Ophthalmol. Vis. Sci*. **35**, 3129 (1994).
13. F. H. Duffy, J. L. Burchfiel, *Brain Res*. **89**, 121 (1975).
14. M. I. Posner, *Q. J. Exp. Psychol*. **32**, 3 (1980).
15. J. A. Brefczynski, E. A. DeYoe, *Nat. Neurosci*. **2**, 370 (1999).
16. C. W. Eriksen, J. D. St James, *Percept. Psychophys*. **40**, 225 (1986).
17. N. G. Muller, O. A. Bartelt, T. H. Donner, A. Villringer, S. A. Brandt, *J. Neurosci*. **23**, 3561 (2003).
18. H. J. Heinze et al., *Percept. Psychophys*. **56**, 42 (1994).
19. S. A. McMains, D. C. Somers, *Neuron* **42**, 677 (2004).
20. C. Morawetz, P. Holz, J. Baudewig, S. Treue, P. Dechent, *Vis. Neurosci*. **24**, 817 (2007).
21. M. M. Muller, P. Malinowski, T. Gruber, S. A. Hillyard, *Nature* **424**, 309 (2003).
22. S. Zeki et al., *J. Neurosci*. **11**, 641 (1991).
23. R. T. Born, D. C. Bradley, *Annu. Rev. Neurosci*. **28**, 157 (2005).
24. S. Treue, J. C. Martinez Trujillo, *Nature* **399**, 575 (1999).
25. Methods are available as supporting material at *Science* Online.
26. P. S. Khayat, H. Spekreijse, P. R. Roelfsema, *J. Neurosci*. **26**, 138 (2006).

27. L. Busse, S. Katzner, S. Treue, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16380 (2008).
28. J. Moran, R. Desimone, *Science* **229**, 782 (1985).
29. K. A. Sundberg, J. F. Mitchell, J. H. Reynolds, *Neuron* **61**, 952 (2009).
30. R. T. Born, *J. Neurophysiol.* **84**, 2658 (2000).
31. We thank Dr. P. Cavanagh and Dr. R. Joober for providing valuable comments and W. Kucharski and S. Nuara for technical assistance. This study was supported by a DAAD fellowship awarded to R.N., CFI and CIHR grants awarded to J.C.M-T, and the Canada Research Chairs Program.

Figure 1, Niebergall et al. 2009

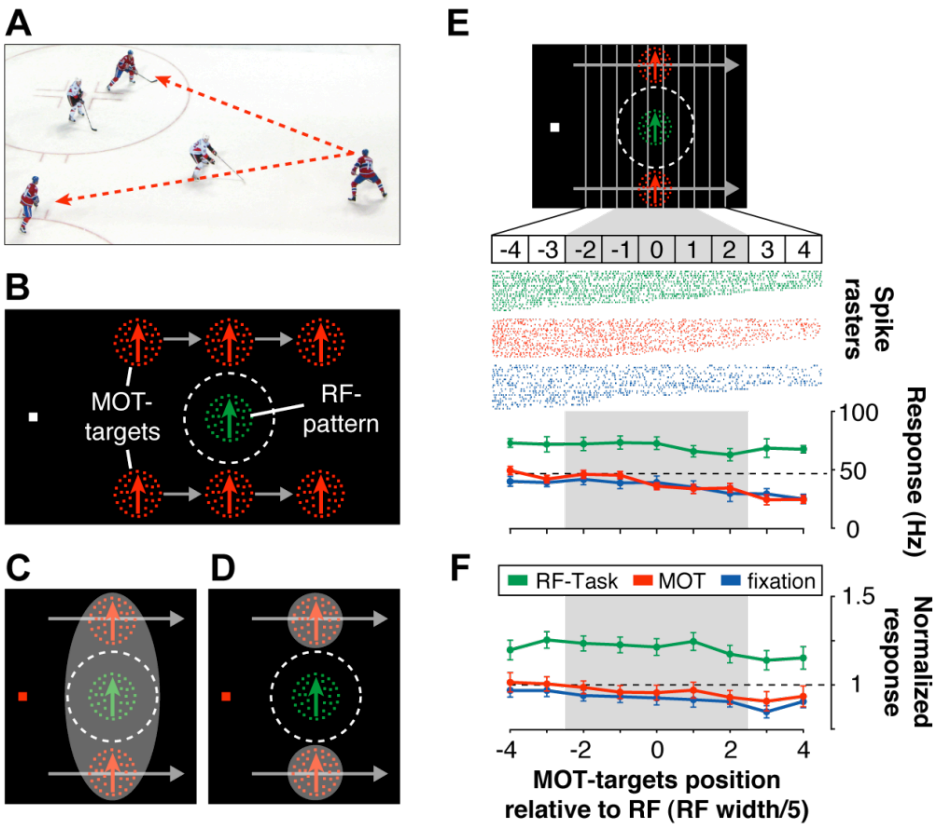


Fig. 1. **(A)** Ice hockey player tracking (arrows) his team members. **(B)** Stimulus configuration. The display illustrates the MOT-targets' (red) trajectories (grey arrows) relative to the fixation spot (white square) and the RF-pattern (green) during a *fixation* trial. The RF-pattern was positioned inside the RF (dashed circle). **(C)** Tracking with a *zoom lens*, or **(D)** with *multifocal attention*. The shaded area represents visuospatial attention. **(E)** Single cell example. Raster plots (*middle*) and average responses (*bottom*) as a function of MOT-targets position (*top*, vertical lines) relative to the RF. Bin sizes for computing the response at each position were defined by '*(RF width)/5*' (see abscissa). The grey shaded area represents the RF-regions. The black dashed line indicates the average response evoked by the RF-pattern in the absence of the MOT-targets. **(F)** Normalized population responses (n=108). For each cell, responses were normalized to the response evoked by the RF-pattern alone (dashed line) and then averaged within each region across neurons. Error bars indicate ±1 s.e.m.

Figure 2, Niebergall et al. 2009

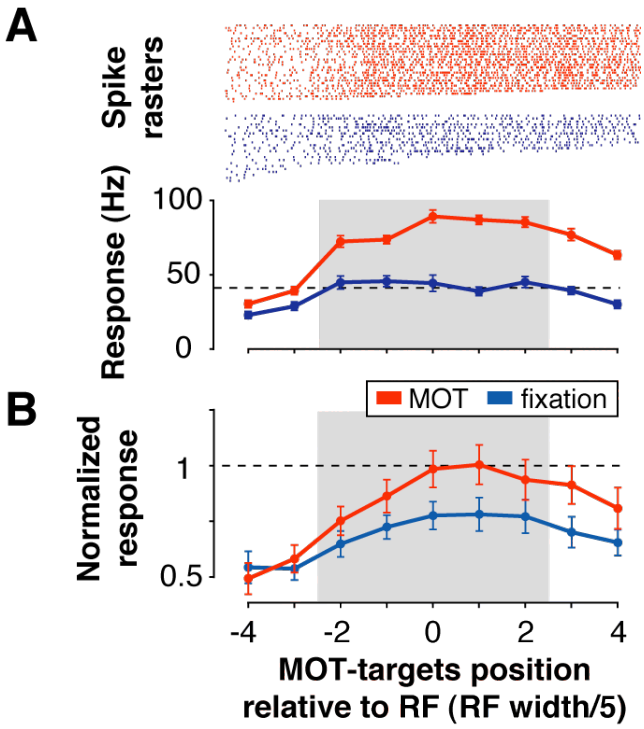


Figure 2. Neuronal responses when the MOT-targets crossed the RF in the absence of the RF-pattern. **(A)** Single cell example responses. **(B)** Normalized population responses (mean ±1 s.e.m) across 42 neurons. The dashed line represents responses to the RF-pattern alone.

Figure 3, Niebergall et al. 2009

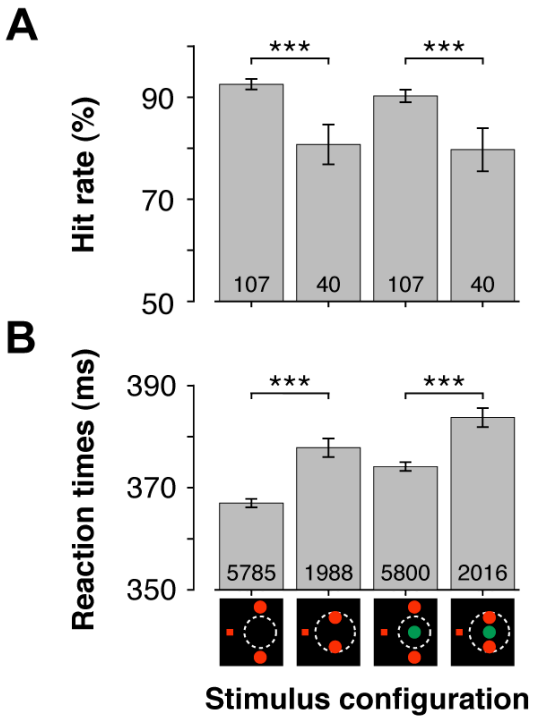


Figure 3. MOT performance for different stimulus configurations. **(A)** Average percentage of correct detections (Hit rate) and (**B)** average reaction times. The numbers on each bar denote the sample size. Error bars display the 95% confidence intervals for the mean. ***$P < 0.0001$ (unpaired *t* test).

Figure 4, Niebergall et al. 2009

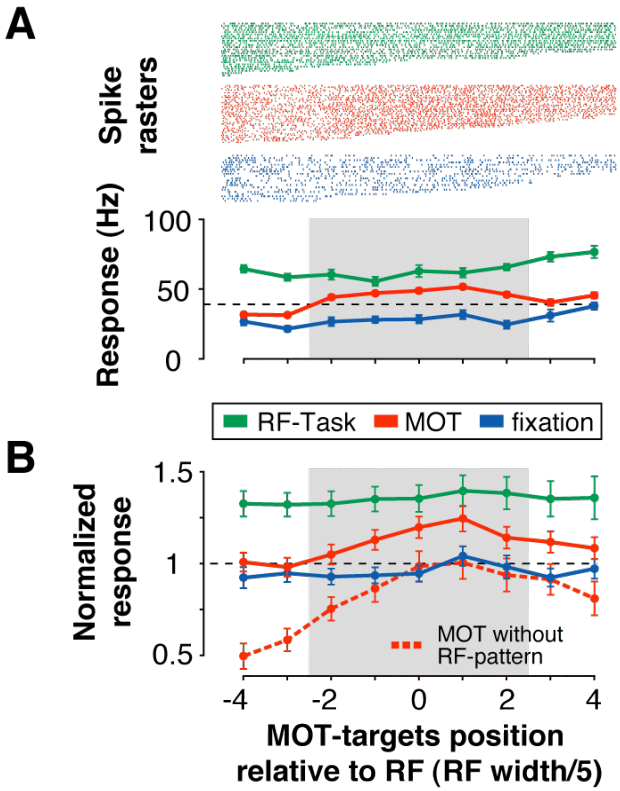


Figure 4. Neuronal responses when the MOT-targets crossed the RF in the presence of the RF-pattern. **(A)** Single cell example responses. (**B)** Normalized population responses (*n*=42). The dashed red line represents responses during *MOT* without the RF-pattern (see Fig. 2). Data expresses mean ±1 s.e.m.

Supporting Online Material (SOM)

Stimuli and task

Stimuli were generated using custom-written software running on an Apple G4 computer and were projected on a rear projection screen using a NEC WT 610 video-projector (NEC Display Solutions of America, Illinois, USA) at a refresh rate of 85 Hz. The animals sat in a primate chair in front of the screen at a viewing distance of 57 cm.

We used moving random dot patterns (RDPs) composed of small dots (one dot area = 0.01 degrees^2) moving behind circular apertures (*1*). The dots could be either green or red (for luminance values see 'Effect of stimulus color or attention to color on cell responses' section). The animals initiated a trial by pressing a button while fixating a white fixation spot (FS, area = 0.06 degrees^2) at the screen's center (**SOM Fig. 1,** leftmost panel). After a delay of 590 ms, three RDPs (positioned at different locations relative to the FS) appeared. Two of the RDPs, the MOT-targets, were presented close to and at the same distance from the FS. The third RDP appeared inside the receptive field (RF) of the recorded neuron (RF-pattern). The dots in the three patterns locally moved in the neuron's preferred direction and at the neuron's preferred speed, which were estimated before starting the recording session (*1, 2*). The MOT-targets started to move at a constant velocity of 3.5 degrees/second following parallel trajectories that passed nearby the stationary RF-pattern without entering the RF boundaries (**SOM Fig. 1,** panels on the right). The distance from each MOT-target to the RF pattern varied along its trajectory, however, it was always similar for both targets. The same principle applied to the distance between the FS and each MOT-target. The dots in the MOT-targets always had the same color (e.g., red), while the RF-pattern dots had a different color (e.g., green). From trial to trial, the colors were randomly interchanged (e.g., MOT-targets red and RF-pattern green, or MOT-targets green and RF-pattern red) to avoid that the animals associated a color with a given stimulus type.

The animals were trained in three different tasks. First, when the color of the FS was white they had to ignore the three RDPs and release the button in response to a change in the spot's luminance (**SOM Fig. 1**, *fixation*, upper panels). When the color of the fixation spot matched the color of the RF-pattern, the animals had to release the button in response to a change in the local speed of the RF-pattern's dots (**SOM Fig.**

1, *RF-task*, middle panels). When the FS color was the same as the MOT-targets' color, the animals had to track the patterns without breaking fixation (see 'Eye position measurements' section) and detect a change in one of the MOT-targets local dot's speed. The change occurred with equal probability (0.5) in either of the MOT-targets (**SOM Fig. 1**, *MOT*, lower panels). The speed change duration was always 110 ms. In order to challenge the animals to maintain attention on the target(s) during the trial, all changes occurred at a random time (820-5060 ms) from trial onset. If the animals released the button within 500 ms after the target change, they received a juice reward. Trials of the different tasks (*fixation, MOT*, and *RF-task*) were randomly intermixed during each recording session.

To make sure that the animals directed attention to the target(s) and ignored the other stimuli, we included approximately 30% of trials containing speed changes in the non-cued RDP(s) (e.g., in the RF-pattern during *MOT*, or in one of the MOT-targets during the *RF-task*). Such changes preceded the target's change, and if the animal released the button in response to them, the trial was aborted and no reward was delivered. In both experiments and for both monkeys the proportion of correct change detections in these trials was above 94% in the *RF-task* and 90% during *MOT*. This indicates that the animals did not respond randomly to any occurring change but selectively responded to changes in the attended target(s).

Performance analysis

The average percentage of correct speed change detections (Hit rate) and average reaction times in the different conditions are summarized in Table 1. *MOT* data were pooled across both MOT-targets.

Table 1. Summary of behavioral data for both monkeys

Exp.	Monkey	Task	Hit rate [%]				#Sessions	Reaction times [ms]				#Trials
			Mean	Median	STD	CI		Mean	Median	STD	CI	
1	Lu	RF	95.1	96.0	4.4	94.0-96.2	64	361.8	360.0	31.0	360.4-363.2	1865
1	Lu	MOT (+RF-pat.)	90.4	92.3	5.9	88.9-91.9	64	378.0	375.0	32.5	377.0-379.0	3717
1	Lu	MOT (-RF-pat.)	93.3	93.9	4.1	92.3-94.3	64	370.4	370.0	30.5	369.4-371.4	3691
1	Lu	Fixation	99.6	100.0	1.3	99.3-99.9	64	304.3	300.0	24.3	303.5-305.1	3192
1	Se	RF	96.7	96.4	5.0	95.1-98.2	43	344.1	345.0	29.6	342.3-345.9	1065
1	Se	MOT (+RF-pat.)	90.0	91.8	7.1	87.8-92.2	43	366.9	365.0	35.7	365.4-368.5	2083
1	Se	MOT (-RF-pat.)	91.4	92.6	6.8	89.3-93.5	43	360.5	360.0	35.5	359.0-362.0	2094
1	Se	Fixation	99.9	100.0	0.2	99.9-100.0	43	318.9	315.0	32.4	317.7-320.1	2672
2	Lu	RF	87.1	95.1	16.3	79.5-94.8	20	372.7	365.0	43.0	368.9-376.5	494
2	Lu	MOT (+RF-pat.)	78.9	83.0	14.8	71.9-85.8	20	380.8	375.0	41.7	378.2-383.4	989
2	Lu	MOT (-RF-pat.)	80.2	84.8	13.7	73.8-86.6	20	374.0	370.0	38.5	371.6-376.4	968
2	Lu	Fixation	99.5	100.0	0.9	99.1-99.9	20	309.7	305.0	32.7	308.0-311.4	1408
2	Se	RF	93.6	95.0	7.3	90.2-97.0	20	384.3	385.0	33.4	381.5-387.1	539
2	Se	MOT (+RF-pat.)	80.6	82.9	11.7	75.1-86.1	20	387.5	380.0	44.5	384.8-390.3	1027
2	Se	MOT (-RF-pat.)	81.3	82.5	10.9	76.2-86.4	20	380.3	375.0	42.3	377.7-382.9	1020
2	Se	Fixation	99.8	100.0	0.6	99.5-100.0	20	338.7	335.0	34.0	337.0-340.4	1474

Exp.1: MOT-targets outside RF

Exp.2: MOT-targets inside RF

STD: standard deviation

CI: confidence interval (95% confidence level)

+RF-pat.: RF-pattern present

-RF-pat.: RF-pattern absent

The following statistical analysis was conducted on the behavioral data shown in **Table 1**. All p-values were obtained from *unpaired* t-tests.

Table 2. Comparisons between conditions. Experiment 1 (MOT-targets outside RF)

Task		Hit rate			Reaction times		
		Monkey			Monkey		
		Lu	Se	Both	Lu	Se	Both
RF	MOT (+RF-pat.)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
RF	MOT (-RF-pat.)	0.0153	0.00011	<0.0001	<0.0001	<0.0001	<0.0001
MOT (+RF-pat.)	MOT (-RF-pat.)	0.0017	0.343	0.0048	<0.0001	<0.0001	<0.0001

Table 3. Comparisons between conditions. Experiment 2 (MOT-targets inside RF)

Task		Hit rate			Reaction times		
		Monkey			Monkey		
		Lu	Se	Both	Lu	Se	Both
RF	MOT (+RF-pat.)	0.1	0.00015	0.00047	0.00052	0.14	0.00075
RF	MOT (-RF-pat.)	0.15	0.00016	0.00097	0.55	0.056	0.71
MOT (+RF-pat.)	MOT (-RF-pat.)	0.77	0.85	0.72	0.0002	0.00016	<0.0001

Table 4. Comparisons between conditions. Experiments 1 and 2

Task exp.1	Task exp.2	Hit rate			Reaction times		
		Monkey			Monkey		
		Lu	Se	Both	Lu	Se	Both
RF	RF	0.0006	0.057	0.00028	<0.0001	<0.0001	<0.0001
MOT (+RF-pat.)	MOT (+RF-pat.)	<0.0001	0.00019	<0.0001	0.026	<0.0001	<0.0001
MOT (+RF-pat.)	MOT (-RF-pat.)	<0.0001	0.00031	<0.0001	0.0011	<0.0001	<0.0001
MOT (-RF-pat.)	MOT (+RF-pat.)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
MOT (-RF-pat.)	MOT (-RF-pat.)	<0.0001	<0.0001	<0.0001	0.0021	<0.0001	<0.0001

In general, both animals performed better during the *RF-task* than during *MOT,* with the exception of monkey Lu when comparing the *RF-task* and *MOT* in experiment 2 (**Table 3**). In experiment 1 we observed a strong modulation of neuronal responses when comparing the *RF-task* vs. *fixation* (~27%), but we found no modulation when comparing *MOT* vs. *fixation* (**Fig. 1f** of the manuscript). It has been demonstrated that the more difficult the task, the stronger the attentional modulation of responses in visual neurons (*3*). Thus, because the most difficult task in our experiments was *MOT* and the easiest was *fixation*, our results cannot be explained by an effect of task difficulty.

We examined whether during *MOT* the animals used alternative tracking strategies rather than simultaneous tracking. First, we tested whether the animals preferentially devoted attention to only one of the MOT-targets. For each session, we quantified the number of correct change detections in each one of the MOT-targets. These data are shown in **SOM Fig. 2a** for both animals. The grey area shows the region of the plot where we considered the animal performed above chance level (> 50% correct detections). In all sessions (n = 107) of experiment 1, both animals did so (**SOM Fig. 2a,** left panels). In these plots, the data concentrate along the unity line suggesting that there were no differences in performance between trials with changes in one or the other MOT-target (P = 0.19, paired t-test for both animals).

In experiment 2 (n = 40), the data followed a similar trend (**SOM Fig. 2a,** right panels). However, the overall performance was lower than in experiment 1, suggesting that the task was more difficult (see also **Fig. 3** of the manuscript). For a small number of experimental sessions (n = 4 for *MOT* with RF-pattern; n = 3 for *MOT* without RF-pattern) we found detection performances below 50% for changes in one of the MOT-targets. Monkey Lu also showed a significant reduction in

performance during *MOT* in the presence of the RF-pattern ($P = 0.02$, paired t-test, **SOM Fig. 2a,** upper right panel).

We also examined reaction times (RTs) to detect changes in the MOT-targets. **SOM Fig. 2b** shows these data for both animals (black and grey) and all experiments. The black and grey dots represent the experimental sessions in which individual comparisons between RTs to changes in each MOT-target were not statistically significant ($P > 0.05$, unpaired t-test). The red and orange points represent the sessions in which we found a significant difference between RTs to changes in one MOT-target relative to the other. For all experiments, the majority of the sessions did not show a significant difference. Also, when comparing the distributions of mean RTs across sessions we did not find significant differences ($P > 0.05$, paired t-test). Thus, we conclude that within sessions the animals did not choose one of the MOT-targets and preferentially attended to it.

Testing the single attentional spotlight switch model of MOT

We tested whether during *MOT* the animals could have rapidly switched a single spotlight of attention from one MOT-target to the other. We reasoned that if this was the case we should see a bimodal distribution of RTs to changes in one of the MOT-targets, because when the change occurred in the attended stimulus, RTs will be shorter than when the change occurred in the unattended stimulus. In the latter scenario, the animal has to switch the spotlight of attention from one stimulus to the other, resulting in slower RTs (*4*). To investigate this possibility, we used the Dip-statistics proposed by Hartigan and Hartigan (*5*) and tested whether RT distributions to changes in one MOT-target were unimodal. The analysis was conducted using *MOT* data in trials where the RF-pattern was present.

SOM Fig. 3a shows one typical example of a RT distribution to changes in one MOT-target. The data appear normally distributed around a single mean and the Dip-statistics did not reveal a deviation from unimodality ($P = 0.1$). In only 9 out of 294 RT distributions (experiment 1: 214; experiment 2: 80) we found P-values lower than 0.05 (**SOM Fig. 3b**). This suggests that 285 distributions were unimodal. A potential problem using the Hartigan's test is that the number of samples in some of the RT distributions was small (< 20), leading to lower power to detect deviations from unimodality. We examined this issue by developing a method that estimated our power to detect deviations from unimodality using the test.

For each experimental session, we divided our sample of RTs, with size n = number of trials, into two groups of size $n/2$ by randomly assigning individual RTs to one or the other group. Thereafter, we added a fixed amount of RT units (i.e., t = 5 ms) to each individual RT of one of the groups and 're-mixed' them. This method produces a shift of t RT units in half of the observations and will, as t becomes larger, necessarily lead to bimodality. We then repeated the Hartigan's test and obtained a P-value. The procedure was repeated after increasing t by 5 ms as many times as needed in order to obtain a P-value < 0.05 in two consecutive iterations. The amount of added RT units in order to obtain statistical significance *(t_s)* was considered as an estimate of our power to detect deviations from unimodality. More importantly, this number represents how fast the animals had to shift attention from one MOT-target to the other in order to account for the observed RT distributions.

SOM Fig. 3c shows the t_s values as a function of sample size for 214 RT distributions. The mean t_s was about 60 ms, suggesting that the animals had to switch attention at a rate of 16 Hz (switches/second) or faster to explain the observed results. Previous studies have demonstrated that monkeys need at least 180 ms (~5.5 Hz) to voluntarily shift attention from one moving RDP to another (*4*). In humans, these numbers are about 150 ms or larger (~6.6 Hz) (*6*). Thus, pooled with the results from **SOM Fig. 2**, these data strongly suggest that in our task, the animals simultaneously tracked both MOT-targets rather than rapidly shifting attention between them.

Eye position measurements

Eye movements were measured using a video-based eye tracking system (Eye Link 1000, SR Research, Ontario, Canada). Horizontal and vertical eye position signals were fed into the stimulus presentation computer that converted them into degrees of visual angle according to a calibration procedure conducted at the beginning of each experimental session. From these signals we computed gaze direction. Monkeys could start a trial if their eye positions were within a 1-degree radius from the FS center. If at any time during a trial gaze position moved outside the fixation window, the trial was aborted without reward.

SOM Fig. 4 shows eye position signals for the different experimental tasks as a function of the MOT-targets' position relative to the RF center averaged across recordings sessions. In animal Lu, we found small systematic deviations toward the side where the RF-pattern was presented, in both, *MOT* and *RF-task*. In animal Se, the

results were similar although deviations were smaller. Given the small size of the deviations (average: 0.24 degrees) relative to the size of the MT neurons RFs (average diameter across recorded cells: 4.7 degrees), and the small size of the RF-pattern relative to the RF size (about 1/3 of the RF size), we consider that these differences in eye position are very unlikely to have caused the pattern of results reported in the manuscript.

Additionally, we found that in experiment 1, neuronal responses during *MOT* and *fixation* were similar, but responses during the *RF-task* were considerably higher relative to the other two tasks. These patterns do not follow the small differences in the average eye position signals seen in **SOM Fig. 4**, which would predict similar differences in response when comparing *MOT* vs. *fixation* and the *RF-task* vs. *fixation*. Thus, we conclude that small variations in eye position cannot cause the pattern of differences and similarities in response amongst the different tasks.

Animal training, preparation, and recordings of single cell responses

After an initial training period, during which the animals became familiar with the primate chair and the experimental setup, we conducted a surgery under general anesthesia and implanted a titanium head post (custom made) in each animal (*1, 7*). After 6 to 8 weeks of recovery, we used the head post to restrain the head and continued the training using eye position measurements. The two animals learned the final task over a period of 6 - 8 months. After that period, we conducted a second surgery and implanted two recording chambers (20 mm diameter, Crist Instruments, Hagerstown, MD, USA) in each animal. The chambers were positioned over the parietal bone using stereotaxic coordinates (*8*) (chamber centers: lateral = ±16 mm, posterior = 5 mm, aligned to inter-aural axis) and were fixed to the skull using dental acrylic and titanium screws. Two weeks after the surgery and once we verified that the animals were completely recovered, we performed an MRI (Siemens 3T Trio MR scanner) in each animal to localize area MT.

During the MRI we used a grid (Crist Instruments, Hagerstown, MD, USA) positioned inside the chamber that provided us with a fixed coordinate system within the chamber. We used glass capillaries (0.3 mm diameter), filled with mineral oil and sealed at the tip with dental acrylic. The capillaries were inserted at fixed positions within the grid (center, 4 mm to the right, left, anterior, and posterior from the center), and were oriented parallel to the chamber walls and orthogonal to the grid surface.

This procedure allowed us to visualize, in the 3D-MRI images, the capillaries orientation relative to the brain anatomical landmarks (**SOM Fig. 5)** and therefore to target area MT during the electrode penetrations using the chamber-fixed coordinate system. In both animals we successfully localized area MT (according to the cells response properties (*2*)) in the first or second electrode penetration after analyzing the MRI data using OsiriX 3.3.2 Imaging software for Apple computers (OsiriX Foundation, Geneva, Switzerland).

Electrode penetrations were conducted using a NAN electrical microdrive and electrode positioning system (Plexon Inc, Texas, USA). During each session, we penetrated the dura mater with a stainless steal guide tube of 500 μm diameter, and then slowly lowered the electrode through the guide tube searching for neural activity. The electrodes were epoxylite insulated with 125 μm shank and impedance between 1 and 5 MΩ (FHC Inc., ME, USA).

During the penetrations, we first identified the different regions where action potentials were isolated through online display of neural signals using Plexon software and recording equipment (Plexon Inc, Texas, USA) and an analogue-digital oscilloscope (Hameg Instruments, MA, USA). We mapped the RF of neurons in each layer of activity using a moving bar and RDPs and determined whether the neurons showed selective responses for motion direction and speed. We usually isolated direction selective responses in the second or third layer of activity (layers were separated by regions of 'no-spiking activity' that we considered as white matter or intra-sulcal space). When a single unit was isolated, we estimated its RF diameter (using a moving bar and RDPs) and its selectivity for spiral motion (expansion, contraction, clockwise-counterclockwise rotation, and their combinations). A neuron was considered as an MT unit if the selectivity for the direction of linear moving RDPs was larger than the selectivity for spiral motion. These selectivities were estimated by online display of the responses to spiral and linear motion stimuli with different average speeds. Additional criteria were that the neuron's RF boundaries did not extend into the opposite hemifield, and that the estimated RF size was not larger than the distance from the FS to the estimated RF center (*1, 2, 6*). Generally we recorded one unit per session, except in three sessions where we recorded two units simultaneously.

RF coordinates analysis

Once the responses of a neuron were isolated, we mapped its' RF with a moving bar and identified the classical RF boundaries (where the preferred stimulus direction ceased to evoke responses), and the putative RF center (where the preferred stimulus direction evoked the strongest response). We then positioned a moving RDP at the RF center and by varying the speed (5 different speeds, 2, 4, 8, 16, 32 degrees/second) and direction (in steps of 30 degrees) in different trial presentations. Thereafter, we obtained an estimate of the cells preferred speed and direction (*1, 2, 6*) and assigned them to the MOT-targets and RF-pattern.

We positioned the RF-pattern at the estimated RF center and the MOT-targets outside at a position where they did not evoke a response from the neuron, even when they were aligned with the RF center (see **Fig. 1e** of the manuscript). We used the formula 'RF diameter ≈ 0.8 * eccentricity' (*9*) to estimate the RF width and divided it into five sub-regions of equal size, oriented orthogonally to the line connecting the FS with the RF center (**Fig. 1a** of the manuscript). The size of one RF region (in space) defined the bin size over which spiking activity was averaged. We extended such regions to both RF sides in order to obtain estimates of cell responses before and after the MOT-targets passed alongside the RF. Our estimates of RF size were conservative in order to avoid that the MOT-targets entered the RF region or its surround.

Nevertheless, we tested whether the MOT-targets had an excitatory influence on neuronal responses along their trajectories. If that were the case, it would indicate that we underestimated the RF width during our mapping procedure and therefore positioned the MOT-targets too close together so that they entered the excitatory region of the RF. **SOM Fig. 6** shows the population responses to the MOT-targets in experiment 1 in the absence of the RF-pattern (data sets with the RF-pattern present are re-plotted for comparison), during *fixation* (purple), and during *MOT* (orange). In both tasks, responses were similar across regions ($P = 0.964$ and $P = 0.916$, one way ANOVA for *fixation* and *MOT*, respectively) and between the two conditions ($P = 0.995$, paired t-test *fixation* vs. *MOT,* pooling data across regions). These results confirm that our RF size estimates were accurate and the translating MOT-targets never entered the RF. They also provide additional evidence in favor of a split of the attentional focus, since it has been demonstrated that attending to the RF in the absence of a visual stimulus, increases neuronal responses (*10*). We did not observe such response increase during *MOT* relative to *fixation*.

In the second experiment, the MOT-targets were positioned close together in a manner that they entered the RF area along their trajectories, crossing it as symmetrically as possible. The data in **Fig. 2** of the manuscript show that our positioning of the stimuli was adequate. We also avoided that the MOT-targets overlapped with the RF-pattern, so they were always separated by a distance larger than the RF-pattern diameter but smaller than the RF diameter.

All the analyzed neural data were obtained from hit trials and truncated at the time of the first stimulus change, whether or not the change happened within the cued stimulus. Because of the variable trial lengths (see **Fig. 1e, 2a, 4a** of the manuscript), a minimum of 5 trials per RF region was required for a cell to be included in the analysis. Responses of each cell across the different RF regions were normalized to the cells average response during *fixation* with only the RF-pattern present. The average response was computed over a time window equivalent to one RF region (RF width/5), following an epoch of 500 ms after stimulus onset. Population responses were obtained by subsequently pooling the normalized responses across neurons in each region.

Effect of stimulus color or attention to color on cell responses

Our task proved to be very difficult for the animals without using colored stimuli. Thus, color was added to make the task easier and the training of the animals feasible. The mean luminance of the different colors were: *red* = 12.8 cd/m^2; *green* = 14.6 cd/m^2. The mean luminance of the background was 0.02 cd/m^2. In a given trial, one of the colors was assigned to the RF-pattern, and the other to the MOT-targets. We reasoned that since MT neurons do not encode stimulus color (*11*), even if the animal used color as a cue, the effects of attention in area MT must have been based on the attended stimulus position (spatial attention).

However, in order to corroborate this assumption, we included two control conditions in our experimental design. In the first one, the animals detected a change in the contrast of the FS and ignored an RDP presented inside the neurons' RF. We included trials in which the RDP moved in the cells' preferred direction and dots could be either red or green. Example trials of this condition for the two different colored RDPs are shown in **SOM Fig. 7a** (left panel). For this example neuron, the responses to both colors were very similar.

We additionally tested whether attention to color could have an effect on MT neurons' response. We instructed the animals to direct attention to the colored stimulus without breaking fixation and detect a speed change that occurred at a random time from the stimulus onset. Data from the same neuron are shown in **SOM Fig. 7a** (right panel). We found a similar response increase for both colors relative to the fixation condition (left panel). But more importantly, when comparing the two attended conditions (attending to the green vs. attending to the red stimulus) responses were similar suggesting that color or attention to color did not cause differences in response.

In order to quantify these observations across the population of neurons we tested 62 units with the two control tasks. We computed color modulation indices (*colorMIs*) between the responses to the two colors during *fixation* and the *RF-task*, using the formula:

$$colorMI = \frac{R_{green} - R_{red}}{R_{green} + R_{red}} \qquad (1)$$

where R_{green} represents the average response to a green RDP inside the RF, and R_{red} the average response to a red RDP. *ColorMI* values higher than zero indicate stronger responses to the green RDP, lower than zero stronger responses to the red RDP, and equal to zero indicate no differences in response between both colored RDPs (*1, 2*).

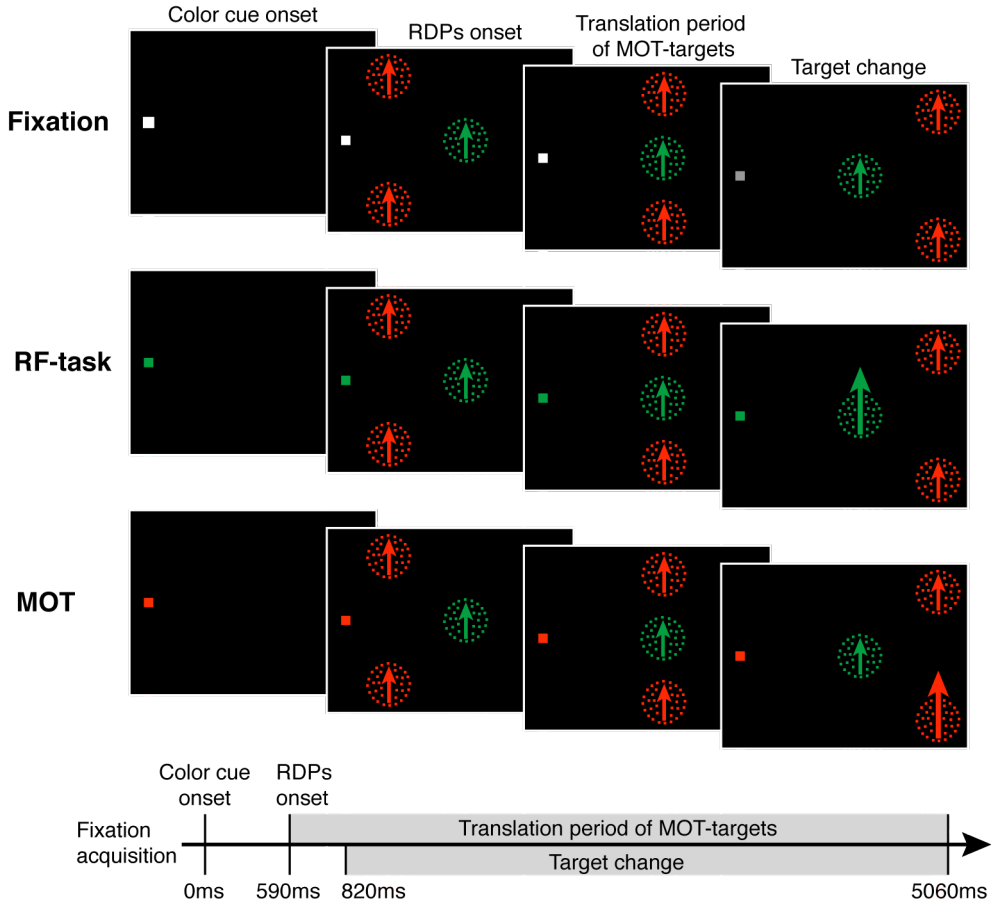
We found that in both cases the *ColorMIs* were not different from zero (**SOM Fig. 7b,** 95% confidence intervals overlap the zero line), suggesting that color or attention to color did not cause the differences in response observed in our experiments. An additional argument in favor of this conclusion is the fact that in experiment 2 we found differences in response between *MOT* and *fixation* only when the MOT-targets entered the RF. Color or attention to color should have produced an effect from trial onset since from the beginning of the trial the stimuli had different colors. We concluded that color may have helped the animals to identify the target but it was not the source of the differences in response we have found in MT neurons.

References

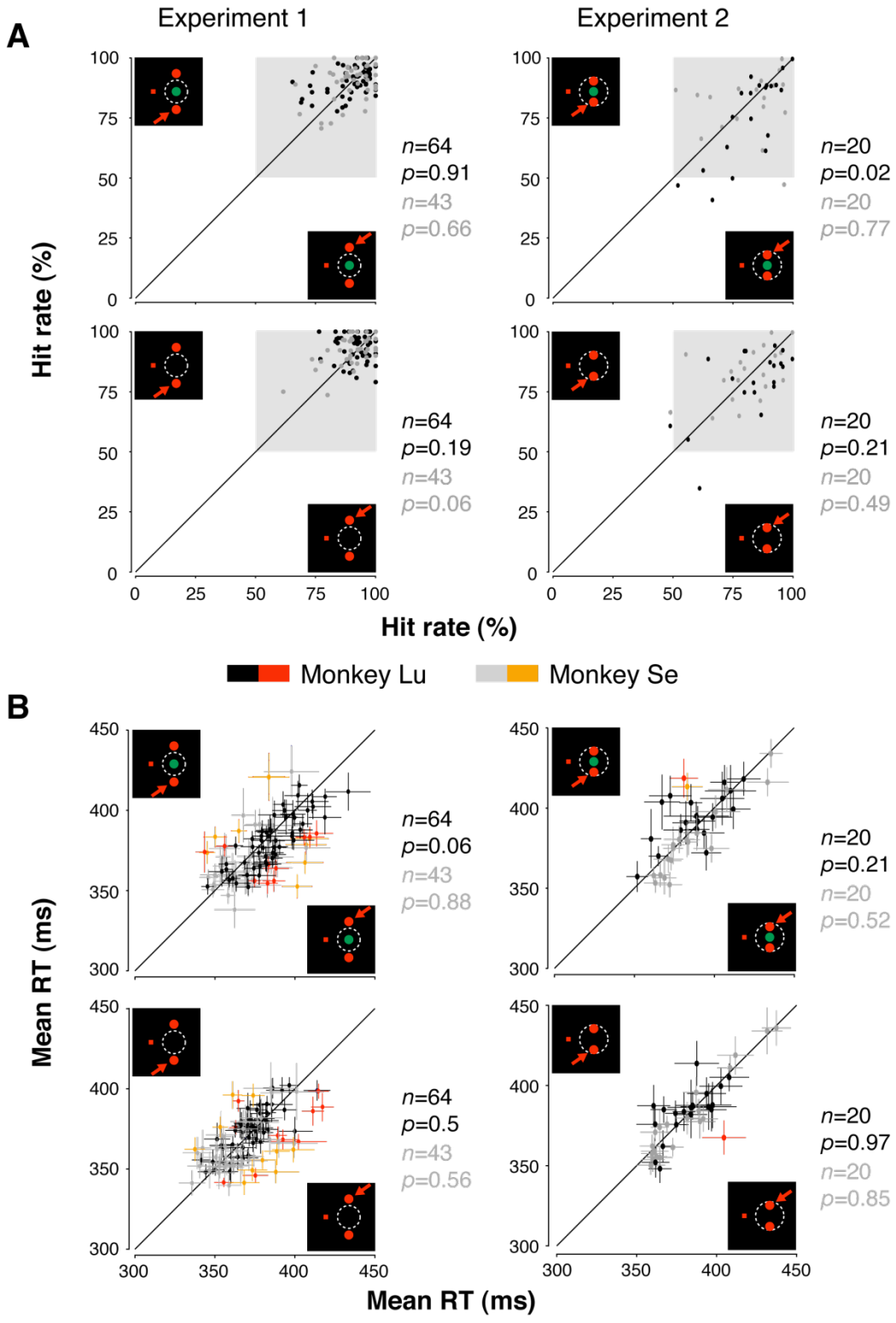
1. J. C. Martinez-Trujillo, S. Treue, *Curr. Biol.* **14**, 744 (2004).
2. S. Treue, J. C. Martinez Trujillo, *Nature* **399**, 575 (1999).
3. H. Spitzer, R. Desimone, J. Moran, *Science* **240**, 338 (1988).
4. L. Busse, S. Katzner, S. Treue, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16380 (2008).
5. J.A. Hartigan, P. M. Hartigan, *The Annals of Statistics* **13**, 70 (1985).
6. P. Cavanagh, G. A. Alvarez, *Trends Cogn. Sci.* **9**, 349 (2005).
7. J. C. Martinez-Trujillo, S. Treue, *Neuron* **35**, 365 (2002).
8. G. Paxinos, X. F. Huang, A. W. Toga, *The Rhesus Monkey Brain in Stereotaxic Coordinates* (Academic Press, San Diego, CA, 2000).
9. K. H. Britten, H. W. Heuer, *J. Neurosci.* **19**, 5074 (1999).
10. S. J. Luck, L. Chelazzi, S. A. Hillyard, R. Desimone, *J. Neurophysiol.* **77**, 24 (1997).
11. A. Thiele, K. R. Dobkins, T. D. Albright, *J. Neurosci.* **19**, 6571 (1999).
12. A. Szucs, *J. Neurosci. Methods* **81**, 159 (1998).

Supporting Online Material figures and legends

Supporting online material, Figure 1



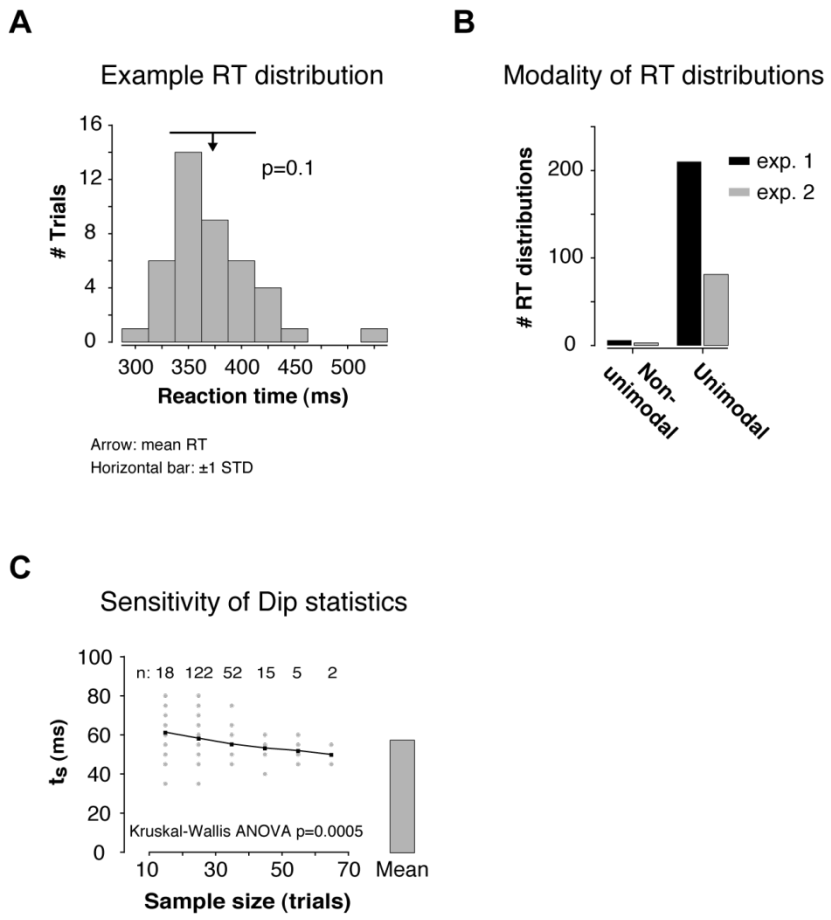
SOM Figure 1. Stimuli and tasks. See text for detailed description of the different tasks. The small square in the leftmost panels depicts the fixation spot (FS), the green RDP the RF-pattern and the red RDPs the MOT-targets. The speed change in the local dots' of the target RDP is indicated by a longer arrow. The timing of events within a trial is shown at the bottom.



SOM Figure 2. MOT-targets were categorized in 'upper' or 'lower' depending on their trajectories' orientation relative to a virtual axis connecting the FS and the RF-pattern center. The upper MOT-target was the stimulus falling above the axis and the

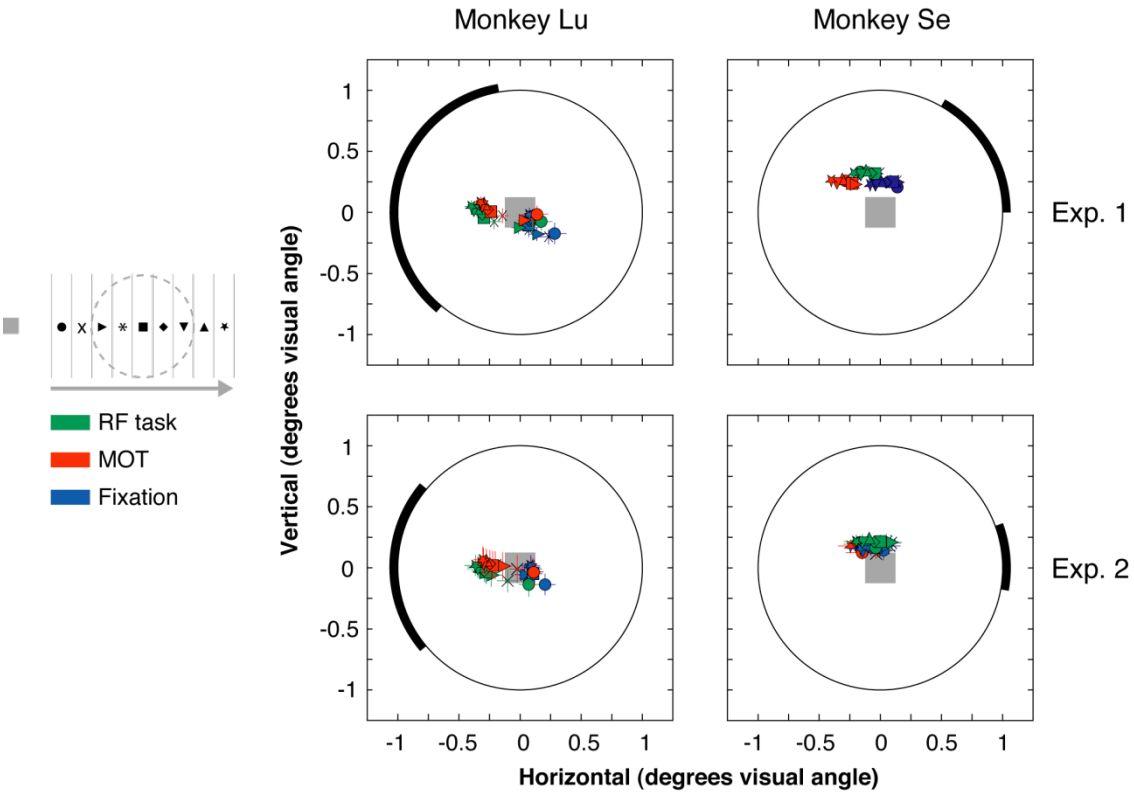
lower MOT-target the stimulus falling below (see inset icons in the graphs; red arrows indicate the pattern containing the speed change). The axis was never vertical. Each data point in the graphs represents data obtained during the same recording session. **(A)** Proportion of correct detections (Hit rate) for changes in each MOT-target during individual recording sessions. The gray shaded area indicates the region were performance was above chance. Data from each monkey are indicated in black and grey, respectively. **(B)** Reaction times (RT) for detecting changes in the 'upper' versus the 'lower' MOT-target across different tasks. Colored data points indicate sessions in which we found significant differences in RT (red: monkey Lu; orange: monkey Se). Error bars represent ±1 s.e.m. Data from experiment 1 (with and without RF-pattern) are shown on the left and from experiment 2 on the right.

Supporting online material, Figure 3



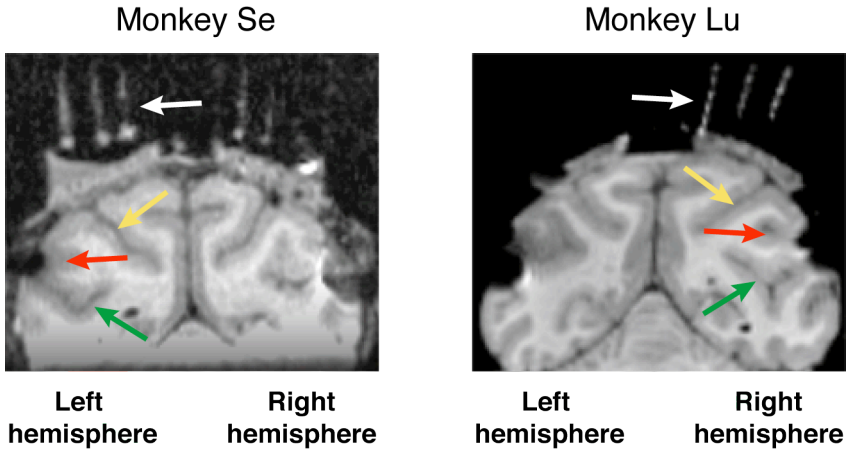
SOM Figure 3. Testing for unimodality of reaction time distributions for speed changes in a MOT-target. (**A)** Example RT distribution during an experimental session. (**B)** Results of the Hartigan's test across the population of RT distributions. **(C)** Sensitivity of the test as a function of sample size. Grey dots represent times t_s (see text) for individual recording sessions. Recording sessions were binned according to their number of trials (*bin width* = 10 trials). The black line connects the mean t_s of each bin and the grey bar on the right shows the mean t_s across all observations. The Kruskall-Wallis ANOVA showed a significant effect of sample size on the sensitivity of the Hartigan's test to detect unimodality. The number of sessions contributing to each bin is indicated. The bar represents the mean t_s across all (n = 214) tested RT distributions.

Supporting online material, Figure 4



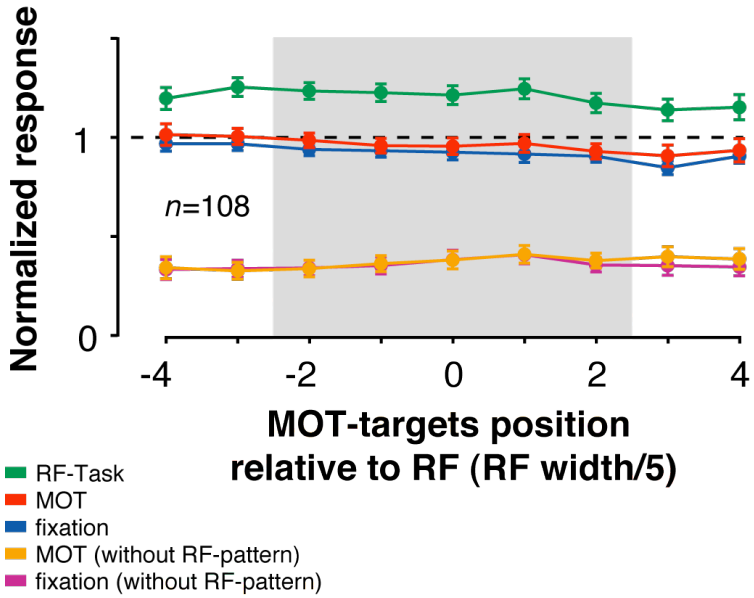
SOM Figure 4. Average eye position signals of each monkey across the different tasks. Individual symbols represent the mean eye position at the time when the MOT-targets crossed each one of the regions (see diagram on the left). The grey square represents the central FS and the circle represents the virtual fixation window constraining the eye position throughout the trial. The thick black lines represent the range of virtual axes (see **SOM Fig. 2** legend) between the fixation spot and the RF-pattern across sessions. Error bars represent ±1 s.e.m.

Supporting online material, Figure 5



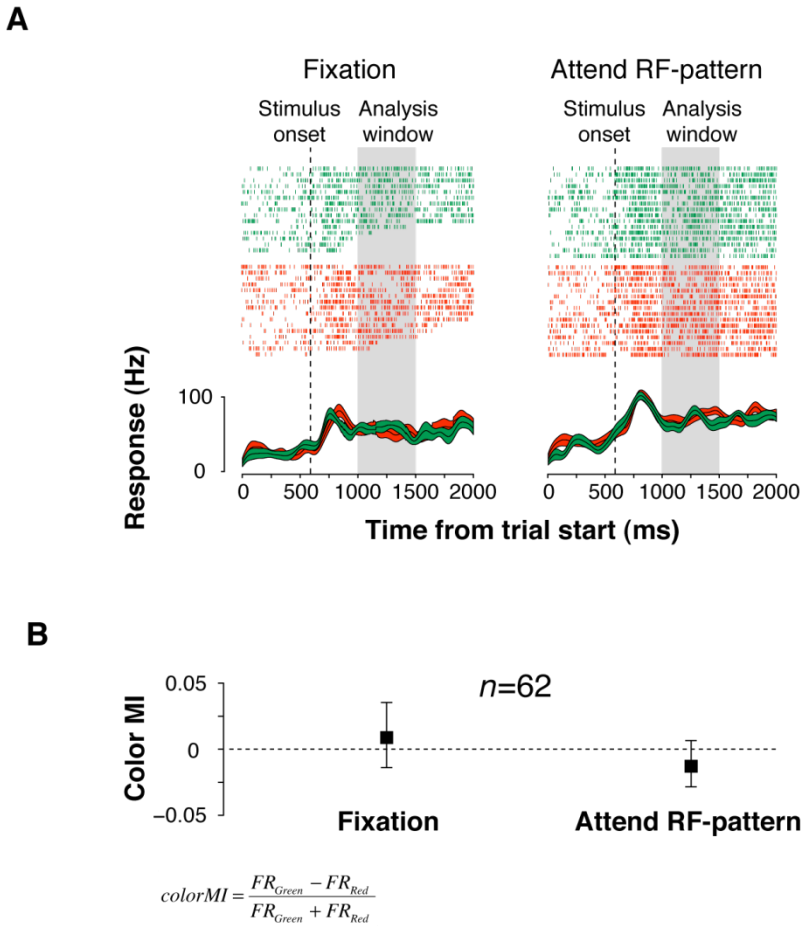
SOM Figure 5. Coronal MRI images and putative recording sites. Anatomical landmarks and fiduciary markers are indicated by the arrows: green, superior temporal sulcus with putative MT recording site; red, lateral fissure; yellow, intraparietal sulcus; white, glass capillaries. Black spots on the brain matter are artifacts produced by the titanium implants positioned on the skull.

Supporting online material, Figure 6



SOM Figure 6. Testing for excitatory surround modulation caused by the translating MOT-targets in experiment 1. Average population responses (±1 s.e.m) obtained during *fixation* (purple) and *MOT* (orange), in the absence of the RF-pattern. For comparison, responses in the tasks where the RF-pattern was present are re-plotted (see Fig. 1f of the manuscript). The grey shaded area represents the RF-regions.

Supporting online material, Figure 7



SOM Figure 7. Effects of stimulus color or attention to color on MT neurons' response. **(A)** Spike rasters and corresponding spike density response function (Gaussian kernel = 20 ms) (*12*) when monkeys ignored (left panel) and attended (right panel) to either a green or a red RDP presented inside the neuron's RF. The upper and lower bounds represent ±1 s.e.m. The grey shaded area represents the time window for which the average firing rate was computed. **(B)** Color modulation indices of the neuronal population. Error bars represent ± 95% confidence intervals.

2.2 Similar perceptual costs for dividing attention between retina- and space-centered targets in humans

Extrafoveal vision critically depends on the ability to direct spatial attention to stimuli positioned at different locations in the visual field. Such locations can be defined in different frames of reference, for example, relative to the fovea (retina-centered), or relative to space (space-centered). One question that has been controversial in studies of attention is whether the allocation of attention preferentially occurs in one of those two frames of reference.

We tested this hypothesis using an experimental paradigm in which human observers focused attention either on a stimulus fixed in retina-centered coordinates or fixed in space-centered coordinates, or divided attention between both stimuli simultaneously. In each experimental condition, perceptual performance was determined by the subject's ability to judge the orientation of a sine wave stimulus. If attention preferentially acts in one frame of reference, then the relative perceptual cost of dividing attention, expressed by a drop in orientation discrimination performance relative to the focused attention conditions, should be lower for targets represented in that reference frame. Contrary to this hypothesis, we found that on average the cost of dividing attention was similar for targets in both reference frames.

This result indicates that the allocation of attention is not significantly biased toward targets represented in one of the tested frames of reference, and suggests that attention is a flexible resource that can be allocated to targets depending on the task demands.

Similar perceptual costs for dividing attention between retina- and space-centered targets in humans

Robert Niebergall[1,2], Lawrence Huang[1], and Julio C. Martinez-Trujillo[1]

[1]*Cognitive Neurophysiology Laboratory, Dept. of Physiology, McGill University, 3655 Promenade Sir William Osler, Montreal, QC, H3G 1Y6, Canada.*

[2]*Cognitive Neuroscience Laboratory, German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany.*

Abstract

Visual-spatial attention has been defined as a spotlight that enhances perceptual performance at attended locations. One subject that has not been fully understood is whether attention is more efficiently allocated to stimuli that remain fixed in one frame of reference (e.g., retina-centered), or whether it could be equally allocated to stimuli fixed in other reference frames. We investigated this issue by asking human observers (n = 13) to covertly attend to sinusoidal gratings that were fixed in different reference frames and to discriminate small changes in their orientation. In the first experiment we quantified orientation discrimination thresholds (ODTs) while subjects pursued a moving dot and either attended to a retina- or a space-centered grating. We then measured ODTs while subjects divided attention between the two gratings. We found that dividing attention proportionally increased ODTs for both target gratings relative to the focused attention condition. In the second experiment, we used the same stimulus configuration and conditions during a fixation task. Here, one grating was retina-, and space-centered, while the other moved in space and on the retina. Again, ODTs during divided attention proportionally increased for both gratings. Moreover, the increases in ODTs were similar to those measured during smooth pursuit (experiment 1). These results show that humans can proportionally divide attention between extrafoveal targets centered in different reference frames and/or with different retinal velocities during both smooth pursuit eye movements and fixations.

Keywords: attention, reference frame, smooth pursuit eye movements, orientation discrimination, retinal velocity

Introduction

Observers can define the position of visual objects in different frames of reference (see Wade & Swanston, 1996 for review), for example, relative to their gaze pointing direction (retina-centered), or relative to space (space-centered). One question that has been targeted in studies of vision is in which frame(s) of reference are attentional re-sources allocated. Some studies have proposed that attention can be efficiently allocated in several reference frames (Barrett, Bradshaw, & Rose, 2003; Behrmann & Tipper, 1999; Danziger, Kingstone, & Ward, 2001; Egly, Driver, & Rafal, 1994; Vecera & Farah, 1994). Others, however, have favored the hypothesis that attention is more efficiently allocated in a retina-centered reference frame (Barrett, Bradshaw, Rose, Everatt, & Simpson, 2001; Golomb, Chun, & Mazer, 2008; Khurana & Kowler, 1987).

Functional imaging studies in humans have reported attentional modulation of brain signals across retinotopically-organized areas of the visual cortex (Brefczynski & DeYoe, 1999; Tootell et al., 1998), suggesting that attention modulates the processing of retina-centered representations. It has also been shown that retina-centered targets lead to faster reaction times and higher accuracy rates during reflexive shifts of attention triggered by exogenous cues (Barrett et al., 2001), as well as during voluntary allocation of attention, induced by endogenous cues (Barrett et al., 2003). The latter study, however, also showed a significant, albeit smaller benefit of allocentric (with respect to other objects in the scene) target representations, suggesting that attention can be effectively directed to non-retina-fixed tar-gets.

Recently, Golomb et al. (2008) showed that after a saccade, identification is faster and more accurate for targets falling on the same retinal location as a pre-saccadic cue, even though the cued location changed relative to space. However, changing task instructions favored the processing of space-centered targets at later time periods after saccades. The authors suggested that the latter effect was due to fast post-saccadic updating of an attentional map in a retina-centered frame, and that the attentional effects triggered by the pre-saccadic cue might be transferred to the new space-centered target, a process resembling predictive remapping (Melcher, 2007). However, this retina-centered attentional advantage isolated by Golomb et al. (2008) may reflect the specific situation of allocating attention during a saccadic eye movement, when active vision is partially suppressed (Duffy & Burchfiel, 1975).

Khurana and Kowler (1987) used smooth pursuit eye movements and asked subjects to pursue a target while simultaneously performing a letter search and discrimination task.

Although these authors did not explicitly test whether attention preferentially acts in a retina- or space-centered frame they found that perceptual judgments about extra-foveal targets moving at the same velocities as the pursued target were more accurate than those about other targets. Complementary to this, Kerzel, Souto and Ziegler (2008) showed reductions in pursuit gain when subjects attended to stationary background objects relative to when they attended to objects moving along with the smooth pursuit target. These results further support the hypothesis that at least under certain circumstances attention is more efficiently allocated to retina-centered targets.

On the other hand, in many situations it may be intuitive to allocate attention in a space-centered reference frame. For instance, attending to a stationary peripheral object during self-motion or smooth pursuit movements requires attention to change position dynamically within a retinotopic map following the retinal displacement of the target representation. In this circumstance, it would be more efficient for attention to act on space-centered representations of the target object, avoiding repetitive shifts of attention across a retinotopic map and leading, at least in theory, to a more robust perceptual enhancement.

The latter hypothesis implies that attention can act on space-invariant representations of objects. Indeed, it has been shown that responses of neurons in regions of the parietal cortex, which show space invariance within their receptive field region (Duhamel, Bremmer, BenHamed, & Graf, 1997; Galletti, Battaglini, & Fattori, 1993), are strongly modulated by attention (Cook & Maunsell, 2002). Some of these parietal cortex neurons seem to encode target motion in space-centered coordinates (Ilg, Schumann, & Thier, 2004; Thier & Ilg, 2005). Moreover, it has been suggested that neurons in human area MT/V5, which responses are also modulated by attention (Saenz, Buracas, & Boynton, 2002), encode the spatial rather than retinal position of visual stimuli (d'Avossa et al., 2007; Tootell et al., 1998; but see Gardner, Merriam, Movshon, & Heeger, 2008).

In addition, studies using multiple-object tracking tasks, have shown that attention can be simultaneously al-located and maintained on several moving objects (for re-view see Cavanagh & Alvarez, 2005). Liu, Austen, Booth, Fisher, Argue and Rempel (2005) showed that to success-fully track multiple objects, subjects relied on the spatial position of objects relative to each other rather than on their retinotopic positions. This result suggests that during tracking attention can be effectively allocated in a space-centered, or in an object-centered frame of reference.

One potential approach to investigate whether attentional resources can be effectively allocated to targets in different frames of reference is by instructing subjects to divide attention between targets centered in the different frames and quantifying changes in performance relative to when attention is focused on each target. This is based on the fact that previous studies have demonstrated that dividing attention between two visual targets both fixed on the retina and in space impairs perception of both stimuli relative to focusing attention on one stimulus at the time (Braun & Julesz, 1998; Joseph, Chun, & Nakayama, 1997; Lee, Itti, Koch, & Braun, 1999; Lee, Koch, & Braun, 1997). In the present study we dissociated the reference frames of two perceptual targets during smooth pursuit eye movements (experiment 1) and fixation (experiment 2) and tested whether dividing attention between them leads to different perceptual costs relative to focusing attention on each target at a time. A lower cost for one target will indicate that attention is more efficiently allocated within that frame. On the other hand, a similar cost for both targets will indicate that attention can be flexibly allocated to targets centered in different frames of reference.

Experiment 1

We tested the ability of human observers to discriminate the direction of a transient change in the orientation of one of two sine wave gratings, eccentrically positioned relative to the gaze pointing direction, while pursuing a moving dot. We measured the subjects' performance in three different conditions: a) when focusing attention on a retina-centered target grating, b) when focusing attention on a space-centered grating, and c) when dividing attention between the two gratings. Importantly, in the three conditions the sensory stimulation was identical but the allocation of attention varied.

METHODS

Subjects

A total of 14 subjects with ages ranging from 20 to 40 years participated in the experiment. All of them had nor-mal or corrected-to normal vision, and signed an informed consent form before starting the experiment. All the procedures used in this study were pre-approved by the ethics committee of the faculty of Medicine at McGill University. Except for two of the authors ('rni', 'jcm'), all subjects were naïve as to the purpose of the experiment.

Apparatus and Stimuli

All stimuli were generated using a custom computer program running on a Macintosh G4 PowerPC. Stimuli were displayed on a CRT monitor (LaCie Inc, Oregon, USA) at a refresh rate of 75 Hz and with a resolution of 1280 x 1024. The stimuli were presented on a white background (*mean luminance* = 41.5 cd/m2). They consisted of two identical sinusoidal gratings with a spatial frequency of 2 cycles per degree and a diameter of 1.6 degrees of visual angle. At the beginning of each trial the two gratings appeared at the center of the screen superimposed on one another (Figure 1A). The smooth pursuit target was a black dot with a diameter of 0.6 degrees of visual angle. Its initial position was either to the left, right, or above the two central sinusoidal gratings at an eccentricity of 8.75 degrees of visual angle.

The experiments were conducted in dim light conditions. Subjects were seated 57 cm away from the screen. Viewing was binocular. In order to avoid subjects using the screen's square edges or any other environmental cue as a reference for perceiving the orientation of the target, they viewed only a portion of the screen through a black cylinder (aperture 45 cm), rendering the visible area circular (Figure 1, see also Ruiz-Ruiz & Martinez-Trujillo, 2008). Responses were collected via a standard keyboard.

Procedure

All stimuli were generated using a custom computer program running on a Macintosh G4 PowerPC. Stimuli were displayed on a CRT monitor (LaCie Inc, Oregon, USA) at a refresh rate of 75 Hz and with a resolution of 1280 x 1024. The stimuli were presented on a white background (*mean luminance* = 41.5 cd/m2). They consisted of two identical sinusoidal gratings with a spatial frequency of 2 cycles per degree and a diameter of 1.6 degrees of visual angle. At the beginning of each trial the two gratings appeared at the center of the screen superimposed on one another (Figure 1A). The smooth pursuit target was a black dot with a diameter of 0.6 degrees of visual angle. Its initial position was either to the left, right, or above the two central sinusoidal gratings at an eccentricity of 8.75 degrees of visual angle.

The experiments were conducted in dim light conditions. Subjects were seated 57 cm away from the screen. Viewing was binocular. In order to avoid subjects using the screen's square edges or any other environmental cue as a reference for perceiving the orientation of the target, they viewed only a portion of the screen through a black cylinder (aperture 45 cm), rendering the visible area circular (Figure 1, see also Ruiz-Ruiz & Martinez-Trujillo, 2008). Responses were collected via a standard keyboard. Target gratings changed their orientation at one of two possible times (after 1170 ms, or after 1730 ms from movement onset), encouraging subjects to attend to the target grating throughout the entire trial. At the end of

the trial (Figure 1A.4) and after the smooth pursuit was completed, subjects reported the direction of the orientation change (left arrow key for counterclockwise or right arrow key for clockwise). Instructions were given to respond as accurately as possible. We did not instruct the subjects to give speeded responses.

Psychophysical data acquisition

We used a weighted staircase method (Kaernbach, 1991) to determine ODTs for each subject and condition. The up(miss)/down(hit) algorithm of 3/1 converged to the orientation change level at 75% correct response rate on the psychometric function. Orientation change intensities were defined in polar angles. They ranged from ±1° (smallest) to ±40° (largest), with positive signs representing clock-wise changes and negative signs representing counterclockwise changes with respect to the vertical. For every possible combination of movement direction of the pursuit spot (from horizontal to vertical meridian and vice versa) and tilt direction (clockwise, counterclockwise) of the target grating, a separate staircase was run, resulting in a counterbalanced number of left and right responses within an experimental block. Each of the staircases was presented in pseudo-randomized order, which prevented the subject from predicting the tilt direction in the current trial. Blocks of each condition (retina-centered and space-centered) consisted of 80 trials each. Staircases within a block were sampled at a fixed number of 20 trials. For each subject, the order of the different blocks was randomized across experimental sessions.

Before entering the data collection sessions, all subjects underwent at least three training sessions of one-hour duration each, after which they reached a stable performance. This avoided the effect of learning as a confounder in our final measurements (Ahissar & Hochstein, 1997), and allowed subjects to become proficient in pursuing the dot without making saccades toward the covertly attended grating(s). During training sessions, auditory feedback on the subject's performance was provided. The feedback was re-moved during experimental sessions. In order to start each staircase at an appropriate difficulty level, close to the potential convergence point, we adjusted the initial orientation change intensity of each staircase to the ODT obtained in the last training session. These values usually differed between targets in the two reference frames. Once the subject entered the experiments the values were kept constant across sessions. Each subject completed four blocks of each condition. In 9 subjects, the smooth pursuit spot moved across the left upper screen quadrant in two of the blocks, and in the other two it moved across the right upper quadrant. In 4 subjects, we tested only two blocks with targets moving in only one direction and located in

the lower right quadrant. The latter was done after analyzing the data from the first 9 subjects and concluding that the quadrant and the direction of the pursuit had no effect on the trial outcome (see auxiliary file Figure 1).

Eye movement recordings

In order to control for smooth pursuit accuracy and potential saccades toward the target grating, eye position signals from both eyes were recorded using an infrared video-based head mounted eye-tracking system (Chronos Vision Inc, Berlin, Germany) with a manufacturer-specified spatial resolution of 0.05°. Head movements were re-strained by using a bite bar, which was customized for each subject using dental impression putty. The eye tracker was fixed to the head and firmly adjusted in order to avoid translational movements of the transducers relative to the subjects' head (DiScenna et al., 1994). We have previously used this procedure in a more demanding task in terms of head stabilization and it proved to be accurate (Ruiz-Ruiz and Martinez-Trujillo, 2008). The horizontal and vertical eye position components were monitored on-line at a sampling frequency of 100 Hz. The eye position data as well as the stimulus parameters and timing of the events during each trial were stored on a computer hard-drive for off-line analysis. The eye tracking software used a pupil-tracking algorithm for converting video images into eye-in-head position signals. Each experimental block started with a calibration procedure in which subjects were instructed to fixate a single dot at four different positions on the monitor (center, 10 degrees up, down, right and left of the center).

DATA ANALYSIS

Smooth pursuit movements

Eye position data were processed and analyzed using Matlab (The Math-Works Inc., Natick, MA, USA). Smooth pursuit data were isolated by visually choosing the onset and offset of the smooth pursuit movement during offline display of trial events and eye movement signals. Eye position signals were filtered by a second-order Butterworth filter with a cut-off frequency of 30 Hz. Blinks were detected by visual inspection of the position signals and were removed from the data.

In order to quantify the measurement error of the smooth pursuit signals and to obtain a threshold velocity criterion for saccade detection we conducted eye movement measurements in three subjects using a different task (see auxiliary files for detailed explanation of the task and results). Small saccades (1 degree visual angle off the smooth pursuit target) reliably

resulted in angular velocities > 20 deg/s (auxiliary files Figure 2D). Based on these measurements, we established that if velocity exceeded 20 deg/s during the two time periods of the target change (1170-1250 ms or 1730–1810 ms after smooth pursuit target motion onset) a trial was determined to contain a saccade and removed from the analysis. We also excluded those trials in which the removed segments comprised more than 30% of the smooth pursuit duration. Across the included subjects, the total number of discarded trials due to saccades was marginal and they were approximately equally distributed across conditions (1% focused retina-centered, 1.1% focused space-centered, 2.8% divided attention; $p = 0.13$, one-way ANOVA). We therefore considered the effect of saccades on ODT measurements negligible. One subject, however, was excluded from the analysis due to the presence of numerous saccades toward the target grating during the orientation change period.

Smooth pursuit velocity

For each trial, the smooth pursuit horizontal and vertical velocity (*vel*) components were computed using Equation 1, where x is the eye position in degrees at time i, and Δt is the time window (50 ms) over which the velocity is calculated.

$$vel_{(i+\frac{\Delta t}{2})} = \frac{x_{(i+\Delta t)} - x_i}{\Delta t} \quad . \qquad (1)$$

The resulting component velocities were then used to compute eye angular velocity (*angVel*) using Equation 2,

$$angVel = \sqrt{vel_{Horizontal}{}^2 + vel_{Vertical}{}^2} \quad . \qquad (2)$$

Angular velocity signals were filtered by a Butterworth filter with a cut-off frequency of 20 Hz (Schütz, Braun, & Gegenfurtner, 2009). Since we were mainly interested in eye velocity signals before and during the time of the orientation change, only the velocity data at the time around the potential target changes (1000 – 1900 ms after smooth pursuit onset) were included in the analysis, leaving out the initial acceleration and compensation of the catch-up saccade. For each trial we computed a smooth pursuit gain by dividing eye angular velocity by the velocity of the smooth pursuit dot. A gain of 1.0 denotes identical eye and target velocities. Mean pursuit gains for each experimental condition were obtained by pooling single trial data.

Smooth pursuit spatial accuracy

In order to measure the smooth pursuit accuracy, we first obtained 2D trajectories by plotting horizontal vs. vertical eye position signals from onset to offset of the smooth pursuit movement for every trial. The distance (*dist*) from each single data point (*i*) to the center of the space-centered target grating (which was also the center of the screen) was computed using Equation 3,

$$dist(i) = \sqrt{x(i)^2 + y(i)^2}\ , \qquad (3)$$

where x is the horizontal eye position (deg), and y is the vertical eye position at sampling point i. In order to obtain an average radius per trial, the mean of the single point distances (radii) was calculated. The mean radii of individual trials were subsequently pooled across experimental conditions and used for statistical analysis. We also estimated the variability of the eye positions around the mean radius by computing, for every trial, the standard deviation (Std) of the distribution of single point distances around the radius. Thereafter, a mean Std was determined for each subject and condition.

Orientation discrimination performance

ODTs were computed using the staircase method (see above). Figure 2 shows staircases from an example subject. Each staircase represents a different combination of reference frame (retina-centered and space-centered) and attentional state (focused attention and divided attention). The orientation change intensity in degrees off the vertical (0°) is plotted as a function of staircase trial number. Staircases started at 6° for the retina-centered trials (white) and at 10° for the space-centered trials (gray) and converged to the threshold level. The smooth pursuit movement direction (upwards) and orientation change direction of the target grating (rightwards) were identical for all staircases.

For each staircase we computed an ODT, defined as the mean orientation change intensity across trials occur-ring after the third reversal point. It represents the magnitude of the orientation change (in degrees) at which the subject correctly discriminates the change direction relative to the vertical in 75% of the trials. In order to ensure that our staircase procedure correctly converged to the ODT, staircases with less than five reversal points were excluded. Across subjects, this led to the dismissal of 10.3% of the total number of recorded

staircases (retina-centered/focused attention: 10.1%; retina-centered/divided attention: 7.5%; space-centered/focused attention: 2.2%; space-centered/divided attention: 11.2%).

To quantify the effects of attentional state (focused vs. divided) and of reference frame (eye-centered vs. space-centered) on ODTs, we computed two indices, an Attentional Modulation Index (AMI) (Treue & Martinez Trujillo, 1999) and a Reference Frame Index (RFI). The AMI was defined by (Equation 4):

$$AMI = \frac{ODT_{dividedAttention} - ODT_{focusedAttention}}{ODT_{dividedAttention} + ODT_{focusedAttention}}. \quad (4)$$

The AMI values vary between -1.0 and +1.0 with negative values representing lower ODTs with divided attention, positive values representing lower ODTs with focused attention and zero representing no differences. Likewise, the RFI was defined by (Equation 5):

$$RFI = \frac{ODT_{space-centered} - ODT_{retina-centered}}{ODT_{space-centered} + ODT_{retina-centered}}. \quad (5)$$

Here, negative values represent lower ODTs in the space-centered condition, positive values represent lower ODTs in the retina-centered condition and zero represents no differences.

Statistical analysis

In order to avoid any unwarranted assumptions about the underlying distributions of ODTs we used non-parametric statistical tests to analyze the performance data. However, the results did not change when parametric tests (t-test) were applied. To facilitate the interpretation of the eye position data and comparison with the existing literature we used parametric statistics.

RESULTS

In order to determine whether subjects accurately pursued the target in the different experimental conditions, we measured the subjects' eye position during trials. We will address both the spatial and the temporal movement profiles during the different conditions.

Smooth pursuit spatial accuracy

We first tested whether eye position trajectories diverged from the pursuit target trajectory.

Figure 3A shows an example of a smooth pursuit trajectory (gray) of one subject superimposed on its mean radius (solid black line). The mean radius represents the average distance of all the data points along the trajectory to the central target. Figure 3B displays the

population mean radii as a function of experimental condition. They are located close to the trajectory of the pursuit dot (dashed line), but were found to be systematically lower ($p < 0.0001$ in all three conditions, t-test). A possible explanation for this result is that subjects tried to minimize the distance between the smooth pursuit dot and the gratings by fixating the edge closer to the target, rather than the dot's center. More importantly, mean radii of the three conditions (*retina-centered* = 8.59, *space-centered* = 8.6, *divided attention* = 8.56) were not significantly different from each other ($p = 0.79$, one-way ANOVA). This indicates that subjects pursued the dot with similar accuracy in all conditions.

In order to test how well each trajectory 'adhered' to the mean radius, we computed the Std of the eye trajectories in each trial. A low Std indicates that the trajectory closely followed the circular motion of the smooth pursuit dot. Figure 3C shows the mean Std across subjects in the three conditions. The average deviations from the mean radius are small (*retina-centered* = 0.233, *space-centered* = 0.221, *divided attention* = 0.227), considering that on aver-age they fall within the size of the smooth pursuit dot (0.6 degrees diameter, Figure 3A). We therefore conclude that the subjects' eye-position trajectories are well described by a quarter of a circle. In addition, the mean Std values of the three groups closely resembled each other ($p = 0.89$, one-way ANOVA), indicating that trajectories did not significantly change across the experimental conditions.

Smooth pursuit velocity

Making accurate smooth pursuit eye movements re-quires the eyes to rotate with similar velocity as the pursuit target. We examined whether subjects showed systematic changes in eye movement velocity relative to the pursuit target velocity across experimental conditions.

Figure 4A shows the means and confidence intervals of velocity profiles in the three experimental conditions for one example subject. Each profile includes trajectories from all valid trials of one experimental block. All profiles show an initial acceleration phase followed by a stable plateau at target velocity (dashed line). During the analysis period used to compute the pursuit gain (highlighted in gray), the velocity profiles closely match the pursuit target velocity, and largely overlap one another.

Figure 4B shows the mean smooth pursuit gain across subjects in the three conditions. The mean values are close to 1 (*retina-centered* = 1, *space-centered* = 0.98, *divided attention* = 0.99, black solid line) and were not significantly different between conditions ($p = 0.82$, one-way ANOVA). These results suggest that subjects tracked the pursuit target with similar

accuracy in the three conditions without making eye movements toward the corresponding target grating.

Orientation discrimination performance

We measured ODTs while subjects covertly attended to either one (focused attention), or both (divided attention) target gratings during the smooth pursuit task. Figure 5A shows the effect of dividing attention on ODTs for targets in the two reference frames. A single data point represents the ODT value during focused attention (abscissa) and its corresponding value during divided attention (ordinate). Each subject contributes with two data points, one representing the retina-centered (white circles) and the other the space-centered (gray triangles) condition.

We first tested whether during focused attention ODTs were different for targets fixed in the two reference frames. The clear separation between ODTs corresponding to retina-centered and space-centered targets (along the abscissa in Figure 5A) suggests consistently higher values for the latter. We quantified this effect by computing RFIs (see Methods) for each subject. The population average RFI (Figure 5B) is significantly above zero (*mean* = 0.34, $p = 0.0002$, Wilcoxon Rank Sign Test), confirming higher ODTs for space-centered relative to retina-centered targets. One likely cause of this difference in ODTs is that besides being centered in different frames, the two targets had different retinal velocity. Indeed, we performed a control experiment in three subjects and found that when increasing the retinal velocity of a target, the ODTs steadily increased (see auxiliary files Figure 3). This indicates that when evaluating possible changes in ODTs in the divided relative to the focused attention conditions, we must compensate for the effect of each target retinal velocity.

The same figure 5A shows that for both targets data points are located above the unity line, suggesting that dividing attention increases ODTs for both retina- and space-centered targets. In order to quantitatively test this result, taking into account the effect of retinal velocity on ODTs, we computed an AMI for each subject (see Methods). This index computes relative changes in ODTs in the divided compared to the focused attention condition for a given target retinal velocity.

Figure 5C displays the average AMI across subjects for both target gratings. Positive values indicate larger ODTs in the divided relative to the focused attention condition. Both average AMIs are larger than zero (*retina-centered* = 0.1 [or 22%]; *space-centered* = 0.08 [or 17%]), indicating that dividing attention increases ODTs (retina-centered: $p = 0.006$; space-centered: $p = 0.001$, Wilcoxon Rank Sign Test). When comparing the mean values

corresponding to both reference frames we found no difference in magnitude of attentional modulation (p = 0.5, Wilcoxon Rank Sign Test for paired data). These data show that the relative cost of dividing attention on ODTs was similar for both the retina-centered and the space-centered target.

DISCUSSION

The results of this experiment demonstrated that a) retina-centered targets lead to lower ODTs than space-centered targets, and b) the relative increase in ODTs caused by dividing attention between both target gratings — each one centered in a different frame — was similar.

We consider a potential explanation for the former result the differences in the targets' retinal velocity. During smooth pursuit the image of the retina-centered target re-mains stationary relative to the retinal surface (Figure 1B, left panel), while the image of the space-centered target constantly changes its position on the retina (Figure 1B, right panel). Changing retinal position over time may pro-duce motion blur (Chung, Levi, & Bedell, 1996; Land, 1999) and deteriorate the quality or stability of the space-centered target representation by orientation-selective visual neurons in areas such as V1 and V4 with retina-fixed receptive fields (Maunsell & Newsome, 1987). This may ultimately lead to impaired perception of orientation changes in this target.

As mentioned in the results, we measured changes in ODTs as a function of retinal velocity in three subjects. Indeed, we found a continuous increase in ODTs with in-creasing retinal velocity (auxiliary files Figure 3). It is also known that performance of extrastriate cortical neurons in area MT of monkeys to encode motion direction improves as a function of the time in which the cell response to a moving stimulus is integrated (Cohen & Newsome, 2009).

Another possibility that may explain the difference in ODTs for the different targets as well as the lack of differences in smooth pursuit gain between different conditions is that smooth pursuit during attentional tracking of the space-centered target grating was more difficult. In order to compensate, subjects may have 'dragged attention away' from the target grating to the pursuit target, keeping pursuit gain similar but causing a decrease in ODTs in the space- relative to the retina-centered target. We will address this issue in experiment 2.

The second result of this experiment was our primary focus of interest. We found a similar increase in ODTs for both targets in the divided relative to the focused attention condition. This result is consistent with previous studies of divided attention between stimuli

that remained fixed in the same frame of reference (Braun & Julesz, 1998; Joseph et al., 1997; Lee et al., 1999; Lee et al., 1997). Our subjects performed on average 20% better in the focused relative to the divided attention condition. The magnitude of this effect is comparable to that reported by Joseph et al. (1997) using Gabor stimuli.

We have previously considered the interaction between smooth pursuit and attention as a potential explanation for the differences in ODTs between the retina- and space-centered targets. This also applies to the comparison between focused vs. divided attention. It is possible that during divided attention subjects devoted more attentional resources to the pursuit target. Dragging attention 'away' from both target gratings might have produced the decreases in ODTs. This explanation is in agreement with studies reporting that smooth pursuit demands attention (Kerzel, Born, & Souto, 2009; Kerzel et al., 2008; Khurana & Kowler, 1987; Lovejoy, Fowler, & Krauzlis, 2009; Made-lain, Krauzlis, & Wallman, 2005; Schutz, Delipetkos, Braun, Kerzel, & Gegenfurtner, 2007; van Donkelaar, 1999; Van Donkelaar & Drew, 2002). Nevertheless, if that were the case, our main conclusion still holds since we obtained a similar drop in ODTs for each target. If available attentional resources were 'dragged away' from the two gratings to be reallocated to the pursuit target, this had to be done proportionally, without favoring either target grating. We will address the possible role of pursuit eye movements in our results in experiment 2.

In summary, our results indicate that subjects did not favor (in their allocation of attentional resources) any of the two target gratings during the divided relative to the focused attention conditions.

Experiment 2

In this experiment, we tested whether our previous results (i.e., the decrease in ODTs in the divided attention condition) were influenced by the fact that we used smooth pursuit eye movements to dissociate the targets' reference frame. As previously mentioned smooth pursuit eye movements require attention. Thus, it is possible that smooth pursuit demanded more attentional resources while attending to the space-centered target and during divided attention, producing increases in the ODTs. This hypothesis would predict that during a fixation task, using the same stimulus configuration and attentional conditions as in experiment 1, the effects caused by divided attention would disappear or be considerably attenuated.

METHODS

Apparatus and Stimuli

Apparatus, viewing conditions, target gratings and smooth pursuit dot (now also the fixation target) were the same as in the previous experiment.

Subjects

A total of seven subjects participated in the experiment. Four subjects (including two of the authors) also participated in experiment 1.

Procedure

In this experiment we measured ODTs while subjects performed a smooth pursuit eye movement similar to the one used in experiment 1 (smooth pursuit condition). In addition, we measured ODTs during a fixation task (fixation condition). The task instructions for the smooth pursuit condition were identical to those of experiment 1.

The only difference was that the pursuit spot always moved across the left upper screen quadrant from the horizontal to the vertical meridian with the target gratings presented in the lower right visual quadrant (Figure 6A). During fixation, subjects maintained gaze on the fixation spot (the pursuit target in the other condition) at the center of the screen (Figure 6B). One of the target gratings remained stationary on the horizontal meridian to the right of the fixation spot while the other moved from the horizontal to the vertical meridian in the lower visual field. Its speed was identical to that of the pursuit target in the smooth pursuit condition. We used this configuration in order to match the retinal locations, trajectory and speed of the target images to those during the smooth pursuit task.

In the fixation condition the definition of reference frames for target gratings is no longer similar to the one during the pursuit condition, since the stationary grating is both retina-centered and space-centered, and the moving grating is neither retina- nor space-centered. We therefore redefined the target stimuli according to their retinal position during a trial as ‘stationary’ and ‘moving’ to provide a common framework for the fixation and pursuit task (Figure 6). The instructions for the attentional task (i.e., focusing attention on the stationary or the moving target or dividing attention between the two) as well as timing of events and eccentricities of the stimuli were identical to those of experiment 1.

Eye-movement Recordings and Analysis

The same methods as previously described were applied for recording and processing of the eye position data and saccade detection. The analysis of the smooth pursuit eye movements

was identical to that of experiment 1. In order to test whether during fixation subjects maintained gaze direction on the central fixation spot, we computed the average eye position for each trial and subsequently pooled those to obtain an average eye position per subject and condition. As a measure of fixation position offset we computed, for each condition, the distance of each subjects' average eye position to the center of the fixation spot.

Orientation Discrimination Performance

We measured ODTs using the staircase method de-scribed in the previous experiment. Staircases with insufficient reversal points (< 5) were discarded (3.4% smooth pursuit; 2.6% fixation). We computed the AMIs for the fixation and smooth pursuit conditions in the same manner as in experiment 1. Additionally, we computed a retinal velocity index (RVI) relating ODTs corresponding to the stationary and moving target in the focused attention condition (Equation 6). For ODTs obtained during smooth pursuit this index is equivalent to the RFI of experiment 1 (see Equation 5).

$$RVI = \frac{ODT_{moving} - ODT_{stationary}}{ODT_{moving} + ODT_{stationary}}. \qquad (6)$$

Negative RVI values represent lower ODTs for the moving target, positive values lower ODTs for the stationary target, and zero no difference.

RESULTS

Eye positions during Smooth Pursuit and Fixation

The results of the analysis of the smooth pursuit eye movements are incorporated in the result section of experiment 1 (Figure 3 and 4), since the experimental conditions were similar. In addition, we tested whether during fixation the subject's eye positions deviated from the central fixation spot.

In all three conditions the average positions fell within the area covered by the fixation spot (gray shaded area). Comparing the offsets across the three conditions revealed no significant difference (p = 0.17, one-way ANOVA), suggesting that the subjects' eye positioning during fixation did not change across conditions. Supporting this finding we also found quasi-homogenous and low saccade detection rates across conditions (*stationary* = 0.7%; *moving* = 1%; *divided attention* = 0.7%).

Discrimination Performance Analysis

We tested whether during fixation dividing attention has a similar effect on ODTs as during smooth pursuit. Figure 8A shows the ODTs of individual subjects for fixation (triangles) and smooth pursuit (circles). The data points are mainly located above the diagonal, suggesting that dividing attention causes ODTs to increase for both moving (gray) and stationary (white) targets. This effect seems to be similar during smooth pursuit (circles), and also during fixation (triangles).

We first quantified the differences in ODTs between the moving and the stationary target within each condition by computing RVIs for both smooth pursuit and fixation (Figure 8B). As anticipated from the previous experiments (Figure 5B and auxiliary Figure 3C), average RVIs were positive, indicating higher ODTs for moving targets. Interestingly, during fixation, target retinal motion seems to have a stronger impact on orientation discrimination performance, leading to a higher average RVI relative to smooth pursuit (*smooth pursuit* = 0.38; *fixation* = 0.65, $p = 0.016$, Wilcoxon Rank Sign Test for paired data).

This effect was mainly due to a decrease in ODTs for the stationary target, and an increase in ODTs for the moving target during fixation relative to during pursuit (i.e., white triangles distributed lower along the diagonal relative to white circles, and gray triangles distributed higher relative to white triangles in Figure 8A). We will refer to a possible explanation for this effect in the discussion.

More importantly, AMIs corresponding to the stationary and moving targets were similar during smooth pursuit (*stationary* = 0.13; *moving* = 0.1, $p = 0.69$, Wilcoxon Rank Sign Test for paired data) and during fixation (*stationary* = 0.13; *moving* = 0.09, $p = 0.69$, Wilcoxon Rank Sign Test for paired data). We further conducted an ANOVA comparing all four AMIs and found no difference between the four groups ($p = 0.77$, Kruskal-Wallis ANOVA, see gray and white bars in figure 8C). This result indicates that dividing attention produced a similar increase in ODTs for the different target types during both fixation and smooth pursuit.

DISCUSSION

This experiment aimed at testing whether the results of experiment 1 were due to the use of smooth pursuit eye movements during the divided and focused attention tasks. We found that dividing attention between a retina-fixed and a retina-moving target resulted in similar impairments of discrimination performance during fixation and during smooth pursuit. These findings discard the hypothesis that interactions between smooth pursuit and the

allocation of attention caused our pattern of results. We will discuss this finding in more detail in the general discussion section.

To our surprise, we found that during fixation the effect of target retinal speed on ODTs was larger than during smooth pursuit. During fixation ODTs were lower for the stationary target and higher for the moving target, relative to those during smooth pursuit (see distribution of data points in figure 8A). At first glance, this is not surprising, since smooth pursuit requires attentional resources, which might have been 'dragged away' from the target gratings to be allocated to the pursuit target. However, this explanation would predict exactly the same effect for the moving target, and we found exactly the opposite.

Although this was not the main focus of our study and clarification of this result needs further investigation, we will elaborate on at least one plausible explanation. It is possible that fixing the target gratings in both retina- and space-centered frames results in lower ODTs relative to fixing them in only one of the reference frames. That would account for the differences between the fixation and pursuit ODTs for the stationary target. Moreover, fixing the target grating in only one frame (e.g., space) may result in lower ODTs than when the target is neither retina- nor space-fixed. During smooth pursuit the moving grating was space fixed but during fixation it was neither retina- nor space-fixed. If one examines figure 8A, the alignment of data point agrees with this hypothesis (stationary: 'retina- and space-fixed' < 'retina-fixed'; moving: 'space-fixed' < 'neither retina- nor space-fixed'). Note that when comparing fixation and smooth pursuit retinal velocity was identical for the stationary targets as well as for the moving targets.

We should state that this hypothesis needs further testing, however, an argument in its favor is that in the primate brain stimulus representations in different frames of references coexist. For example, in early visual cortex representations are retinotopically organized (Gardner et al., 2008) while in higher areas of the parietal and frontal cortex representations are more space-invariant (Martinez-Trujillo, Medendorp, Wang & Crawford, 2004; Olson & Gettner, 1995). This space invariance must arise from the early retinotopic representations through neural computations. One could easily imagine that when the two reference frames coincide, computations are less demanding. However, when the representation is not fixed in either frame, computations are the most demanding since the only available option is updating in one or the other frame while the eccentrically positioned target moves on the retina and/or in space (Merriam, Genovese & Colby, 2003). The latter operation may be the most computationally demanding since it may require re-mapping of receptive fields.

General Discussion

The main contribution of this study was to demonstrate that attention can be proportionally divided between targets in different frames of reference. This effect was not dependent on the attended targets' retinal velocity, or on whether subjects were pursuing a dot or fixating it when covertly attending to the targets.

Our results apparently contradict those reported by Golomb et al. (2008). These authors found that during sac-cades attention is preferentially allocated to retina-centered targets, and that allocation to targets in other frames of reference is due to re-mapping of the focus of attention in retinal coordinates. A potential explanation for this apparent discrepancy is that the processes underlying the allocation of attention during saccades might be different from those during smooth pursuit. One feature of saccades that makes them 'special' is saccadic suppression, i.e. stimulus representations in visual neurons are partially suppressed during a saccade (Duffy & Burchfiel, 1975). Thus, it seems plausible that attention remains attached to the retinal position where the target was located prior to the saccade. Interestingly, Golomb et al. (2008) reported that the advantage for targets in a retina-centered frame evanished when more time was available to perform the task: subjects could then equally allocate attention to space-fixed targets, as was also the case in our study.

Supporting the previous idea, Horowitz, Holcombe, Wolfe, Arsenio & DiMase (2004) have made a distinction between the time needed to make saccade-like shifts of attention (attentional saccades) and smooth pursuit-like shifts of attention (attentional pursuit). It is possible that these two types of attentional shifts differ in their mechanisms. While attentional saccades may favor retina-centered tar-gets, attentional pursuit may be equally effective for both, retina- and space-centered targets. In our task the likely mechanism of attentional shifts was attentional pursuit. That may explain why attention did not favor one or the other target. Other studies have also supported the notion that attention can be allocated to targets in various reference frames (Barrett et al., 2003; Behrmann & Tipper, 1999; Danziger et al., 2001; Egly et al., 1994; Vecera & Farah, 1994).

Finally, another plausible hypothesis is that when attention is automatically (involuntarily or exogenously) allocated to a target, the preferred frame of reference is retina-centered, while voluntary or endogenous attention can be equally allocated in different reference frames. This notion may be supported by the view that early visual areas encode the position of stimuli in retina-centered coordinates (Gardner et al., 2008), while space- or object-centered representations do not appear in the visual system until later stages after

visual signals have undergone processing in early areas (Martinez-Trujillo et al., 2004; Olson & Gettner, 1995). If one considers that exogenous attention is mostly a bottom-up driven process and the saliency of the stimulus is primarily encoded in retinal-centered maps in visual areas one may anticipate retina-centered allocation of attention. On the other hand, because endogenous attention is a top down process, it may have access to multiple representations at different levels in the hierarchy of processing.

Conclusions

We conclude that humans can proportionally divide attention between targets in different frames of reference, and or with different retinal velocities, during both smooth pursuit and fixation. This demonstrates that visual attention is a flexible mechanism that modulates information processing in the human brain by accessing stimulus representations at different levels in the visual hierarchy.

Acknowledgments

We would like to thank Walter Kurcharski for technical assistance and Navid Sadeghi Ghandehari for helpful comments on the manuscript.

This study was supported by a DAAD fellowship awarded to R.N., CFI and CIHR grants awarded to J.C.M-T, and the Canada Research Chairs Program.

Commercial relationships: none.

Corresponding author: Julio C. Martinez-Trujillo.

Email: julio.martinez@mcgill.ca

Address: Cognitive Neurophysiology Laboratory, Dept. of Physiology, McGill University, 3655 Promenade Sir William Osler, Montreal, QC, H3G 1Y6, Canada.

References

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387,* 401-406.

Barrett, D. J., Bradshaw, M. F., & Rose, D. (2003). Endogenous shifts of covert attention operate within multiple coordinate frames: evidence from a feature-priming task. *Perception, 32,* 41-52.

Barrett, D. J., Bradshaw, M. F., Rose, D., Everatt, J., & Simpson, P. J. (2001). Reflexive shifts of covert attention operate in an egocentric coordinate frame. *Perception, 30,* 1083-1091.

Behrmann, M., & Tipper, S. P. (1999). Attention accesses multiple reference frames: evidence from visual neglect. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 83-101.

Braun, J., & Julesz, B. (1998). Withdrawing attention at little or no cost: detection and discrimination tasks. *Perception & Psychophysics, 60*, 1-23.

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience, 2*, 370-374.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*, 349-354.

Chung, S. T., Levi, D. M., & Bedell, H. E. (1996). Vernier in motion: what accounts for the threshold elevation? *Vision Research, 36*, 2395-2410.

Cohen, M. R., & Newsome, W. T. (2009). Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *Journal of Neuroscience, 29,* 6635-6648.

Cook, E. P., & Maunsell, J. H. (2002). Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nature Neuroscience, 5,* 985-994.

d'Avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D. C., & Morrone, M. C. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nature Neuroscience, 10*, 249-255.

Danziger, S., Kingstone, A., & Ward, R. (2001). Environmentally defined frames of reference: their time course and sensitivity to spatial cues and attention. *Journal of Experimental Psychology: Human Perception & Performance, 27*, 494-503.

DiScenna A. O., Das, V., Zivotofsky, A. Z., Seidman, S. H., & Leigh, R. J. (1995). Evaluation of a video tracking device for measurement of horizontal and vertical eye rotations during locomotion. *Journal of Neuroscience Methods, 58*, 89-94.

Duhamel, J. R., Bremmer, F., BenHamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature, 389*, 845-848.

Duffy, F.H., & Burchfiel, J.L. (1975). Eye movement-related inhibition of primate visual neurons. *Brain Research, 89,* 121-132.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*, 161-177.

Galletti, C., Battaglini, P. P., & Fattori, P. (1993). Parietal neurons encoding spatial locations in craniotopic coordinates. *Experimental Brain Research, 96*, 221-229.

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *Journal of Neuroscience, 28*, 3988-3999.

Golomb, J. D., Chun, M. M., & Mazer, J. A. (2008). The native coordinate system of spatial attention is retinotopic. *Journal of Neuroscience, 28*, 10654-10662.

Horowitz, T. S., Holcombe, A. O., Wolfe, J. M., Arsenio, H. C., & DiMase, J. S. (2004). Attentional pursuit is faster than attentional saccade. *Journal of Vision, 4,* 585-603.

Ilg, U.J., Schumann, S., & Thier, P. (2004). Posterior parietal cortex neurons encode target motion in world-centered coordinates. *Neuron, 43,* 145-151.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a 'preattentive' feature search task. *Nature, 387*, 805-807.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics, 49*, 227-229.

Kerzel, D., Born, S., & Souto, D. (2009). Smooth pursuit eye movements and perception share target selection, but only some central resources. *Behavioural Brain Research* 201, 66-73.

Kerzel, D., Souto, D., & Ziegler, N. E. (2008). Effects of attention shifts to stationary objects during steady-state smooth pursuit eye movements. *Vision Research, 48*, 958-969.

Khurana, B., & Kowler, E. (1987). Shared attentional control of smooth eye movement and perception. *Vision Research, 27*, 1603-1618.

Land, M. F. (1999). Motion and vision: why animals move their eyes. *Journal of Comparative Physiology [A], 185*, 341-352.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience, 2*, 375-381.

Lee, D. K., Koch, C., & Braun, J. (1997). Spatial vision thresholds in the near absence of attention. *Vision Research, 37*, 2409-2418.

Liu, G., Austen, E. L., Booth, K. S., Fisher, B. D., Argue, R., Rempel, M. I., et al. (2005). Multiple-object tracking is based on scene, not retinal, coordinates. *Journal of Experimental Psychology: Human Perception & Performance, 31*, 235-247.

Lovejoy, L. P., Fowler, G. A., & Krauzlis, R. J. (2009). Spatial allocation of attention during smooth pursuit eye movements. *Vision Research, 49*, 1275-1285.

Madelain, L., Krauzlis, R. J., & Wallman, J. (2005). Spatial deployment of attention influences both saccadic and pursuit tracking. *Vision Research, 45*, 2685-2703.

Martinez-Trujillo, J. C., Medendorp, W. P., Wang, H., & Crawford, J. D. (2004). Frames of reference for eye-head gaze commands in primate supplementary eye fields. *Neuron, 44*, 1057-1066.

Maunsell, J. H., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10,* 363-401.

Melcher, D. (2007). Predictive remapping of visual features precedes saccadic eye movements. *Nature Neuroscience, 10*, 903-907.

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2003). Spatial updating in human parietal cortex. *Neuron, 39*, 361-373.

Olson, C. R., & Gettner, S. N. (1995). Object-centered direction selectivity in the macaque supplementary eye field. *Science, 269*, 985–988.

Ruiz-Ruiz, M., & Martinez-Trujillo, J. C. (2008). Human updating of visual motion direction during head rotations. *Journal of Neurophysiology, 99*, 2558-2576.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002) Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5,* 631-632.

Schütz, A. C., Braun, D. I., & Gegenfurtner K. R. (2009). Object recognition during foveating eye movements. *Vision Research*. 49, 2241-2253.

Schütz, A. C., Delipetkos, E., Braun, D. I., Kerzel, D., & Gegenfurtner, K. R. (2007). Temporal contrast sensitivity during smooth pursuit eye movements. *Journal of Vision, 7*, 3 1-15.

Thier, P., & Ilg U. J. (2005). The neural basis of smooth-pursuit eye movements. *Current Opinion in Neurobiology, 15,* 645-652.

Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron, 21*, 1409-1422.

Treue, S., & Martinez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399*, 575-579.

Van Donkelaar, P. (1999). Spatiotemporal modulation of attention during smooth pursuit eye movements. *Neuroreport, 10*, 2523-2526.

Van Donkelaar, P., & Drew, A. S. (2002). The allocation of attention during smooth pursuit eye movements. *Progress in Brain Research, 140*, 267-277.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123*, 146-160.

Wade, N. J., & Swanston, M. T. (1996). A general model for the perception of space and motion. *Perception, 25*, 187-194.

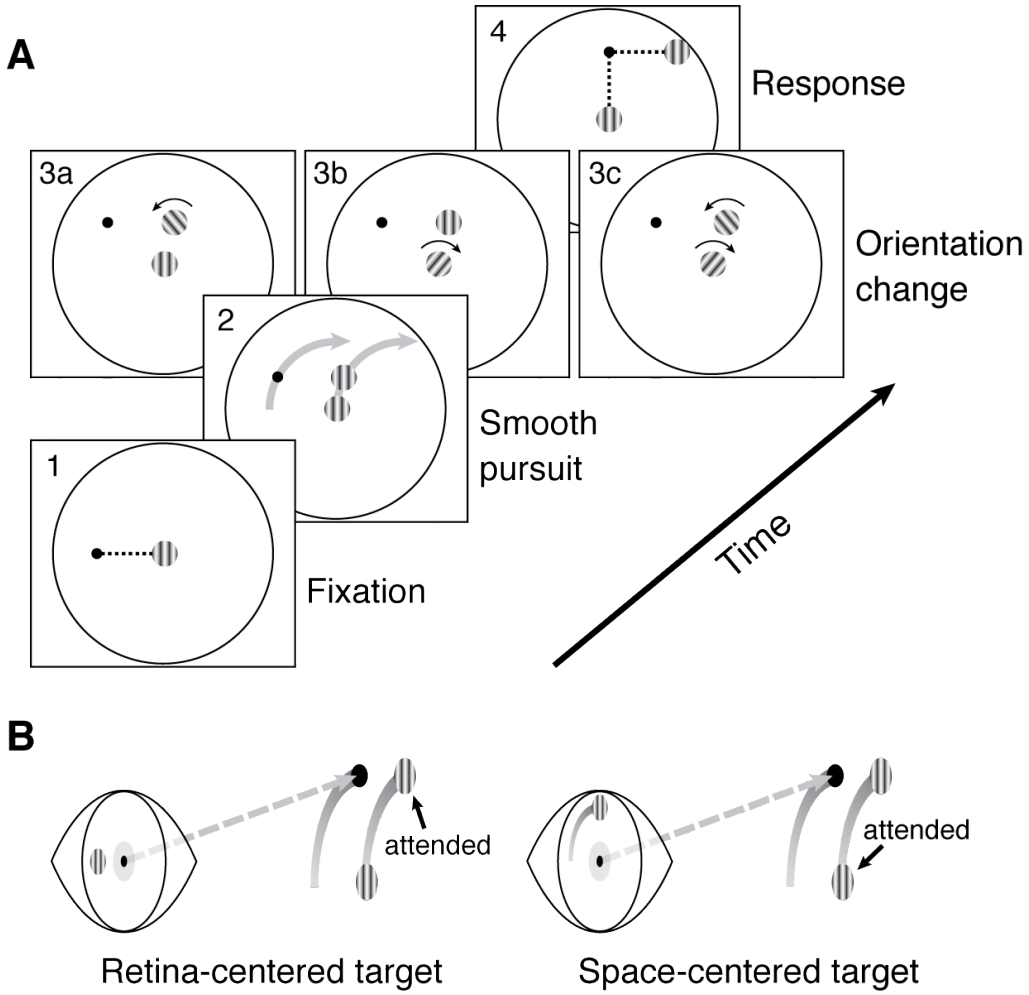


Figure 1. Experiment 1. A) Trial sequence. The panels represent the visual display, the black dot the pursuit target and the gratings the orientation discrimination targets. The gray and black curve arrows represent the grating's trajectory, the pursued target trajectory, and the grating tilt (clockwise and counter-clockwise) respectively. Dotted lines indicate stimulus eccentricity. 3a,b,c illustrate the different attentional conditions. Attention was focused on: a) the retina-centered target or b) the space-centered target or c) was divided between both. B) Hypothetical retinal image of the attended gratings. The left cartoon represents a behind view of the eye and retinas, and the right one the stimulus display with the pursued dot and the gratings' trajectory. The gray dashed arrows represent gaze direction.

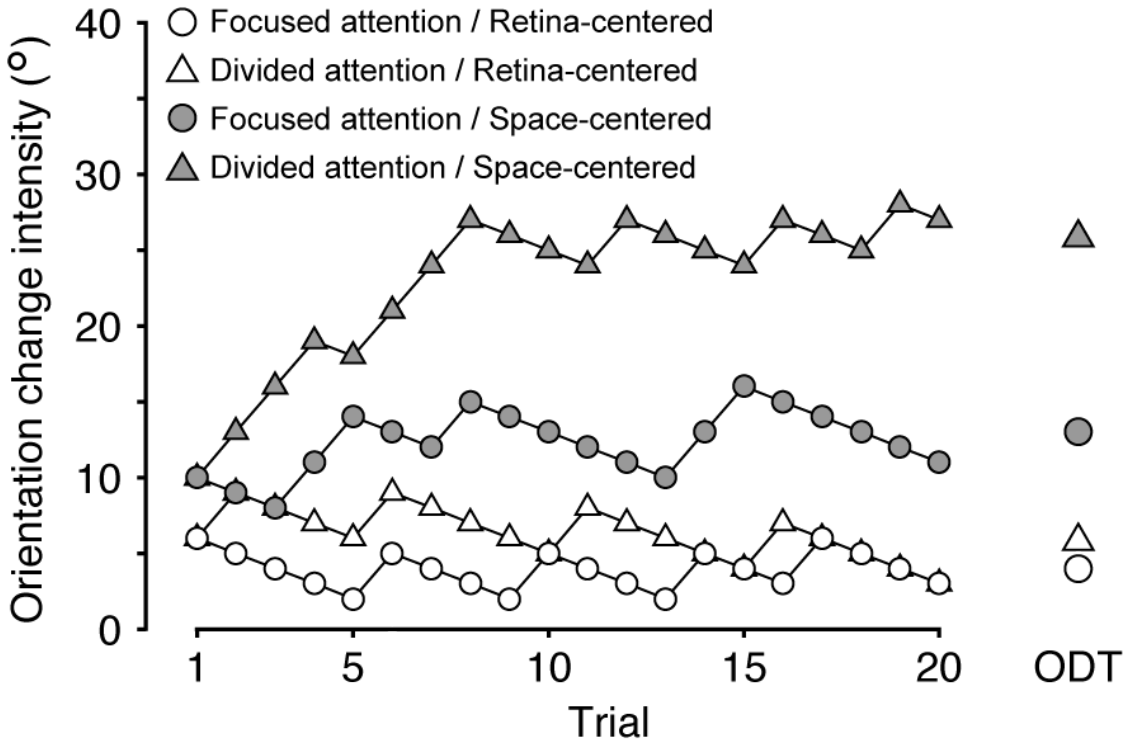


Figure 2. Experiment 1. Staircases corresponding to one example subject (’jcm’). The abscissa represents the trial number and the ordinate the orientation change intensity. The data points on the right are the ODTs, obtained by averaging the orientation change intensities corresponding to each staircase after the third reversal point.

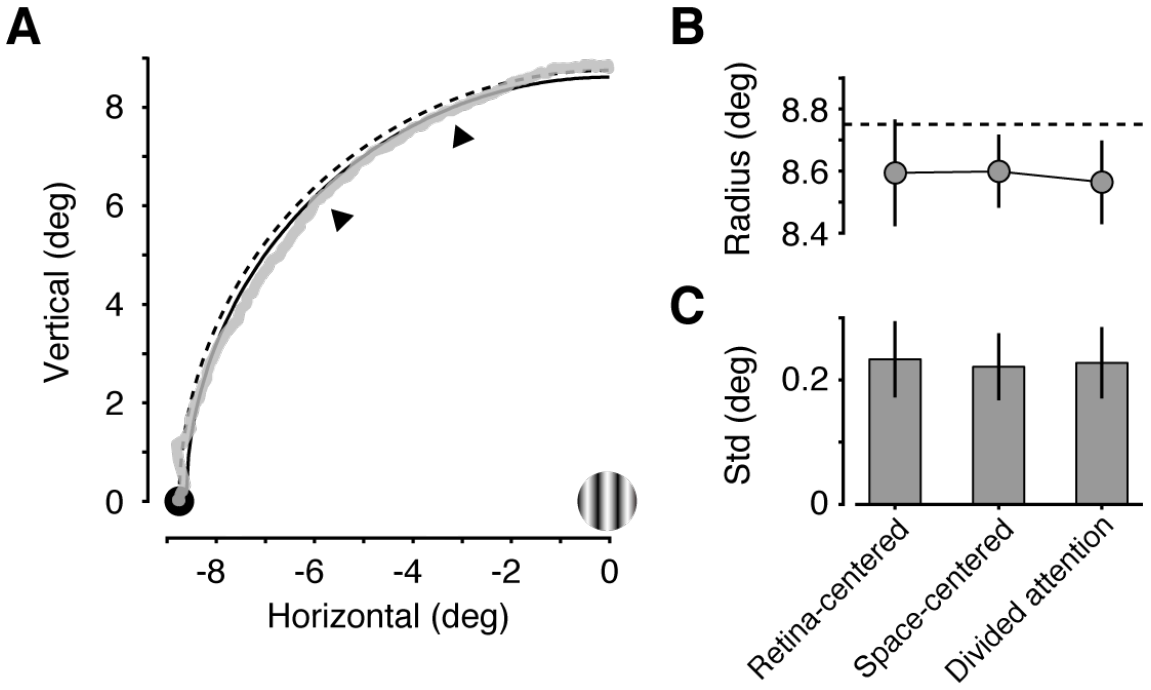


Figure 3. Experiment 1. Smooth pursuit accuracy. A) Example trial. The gray trace depicts horizontal and vertical positions of a smooth pursuit eye movement trajectory moving from left bottom to top right, The black dashed line represents the trajectory of the pursuit dot (black disc) and the black solid line the mean radius corresponding to the smooth pursuit movement trajectory. The black arrows represent the positions at which the gratings could change orientation. The bottom right grating represents the space-centered target. B) Mean radii values (± 1Std) of pursuit trajectories in the different experimental conditions across subjects (*n* = 13). The dashed line represents the trajectory of the pursuit dot. C) Mean standard deviation of the radii (± 1Std).

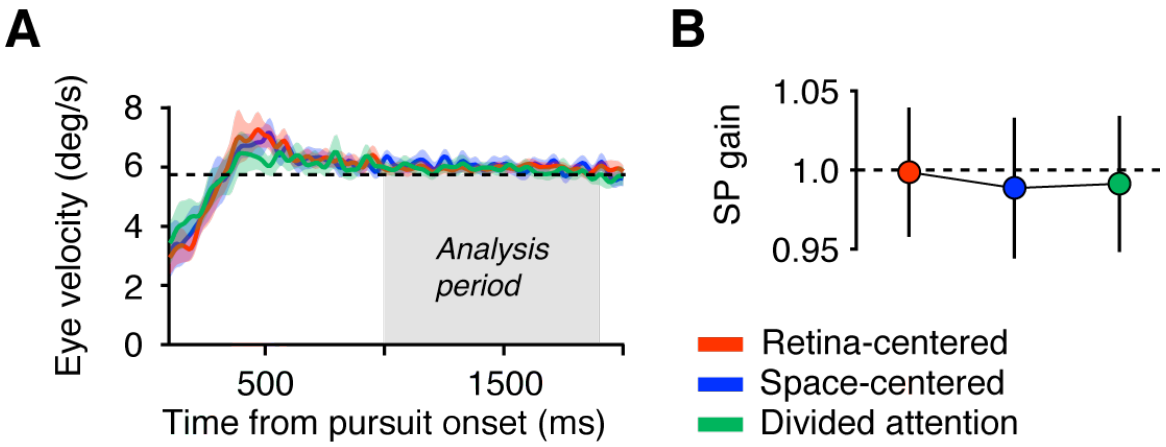


Figure 4. Experiment 1. A) Mean eye angular velocities as a function of trial time in the different conditions for one subject. Eye velocities (colored lines) are superimposed on the velocity of the smooth pursuit spot (dashed line). The shaded areas represent the 95% confidence intervals of the mean. B) Average smooth pursuit (SP) gain across subjects for the different conditions (n = 13). For each subject averages were computed over the time period indicated by the gray shaded area in 'A'. Error bars indicate Std.

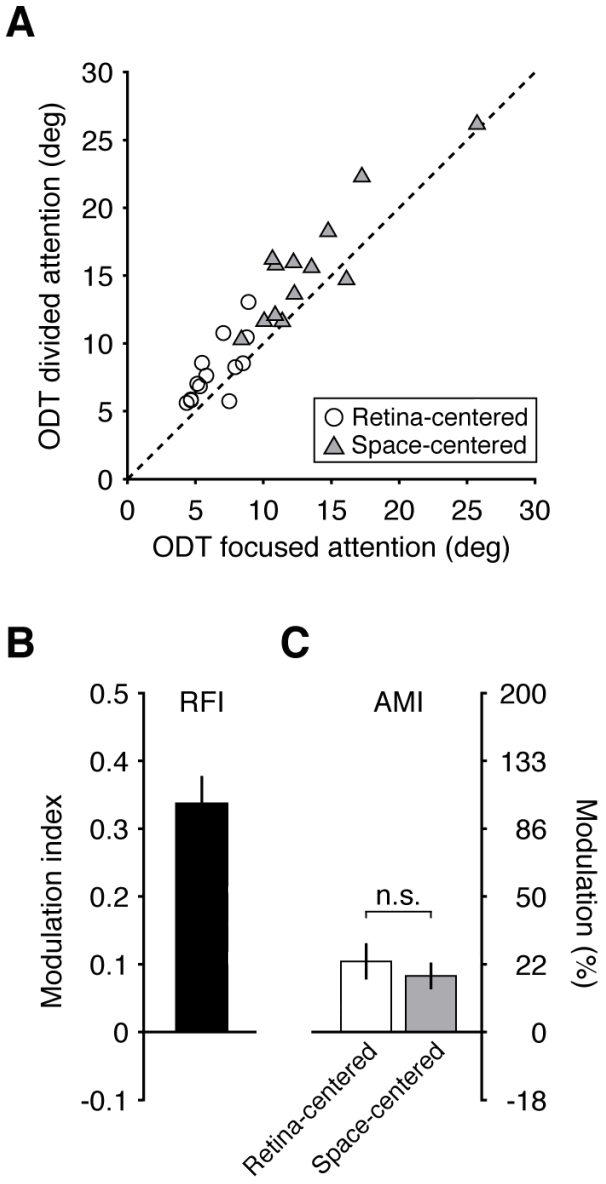


Figure 5. Experiment 1. A) Effect of dividing attention between targets in retina-centered and space-centered reference frames. Each symbol represents the average ODT of an individual subject. ODTs for retina-centered (white circles) and space-centered targets (gray triangles) during focused attention are plotted against their corresponding ODTs during divided attention (*n* = 13). B) Average reference frame index (RFI) computed on focused attention data in 'A'. C) Average attentional modulation index (AMI) as a function of reference frame. The error bars represent SEM.

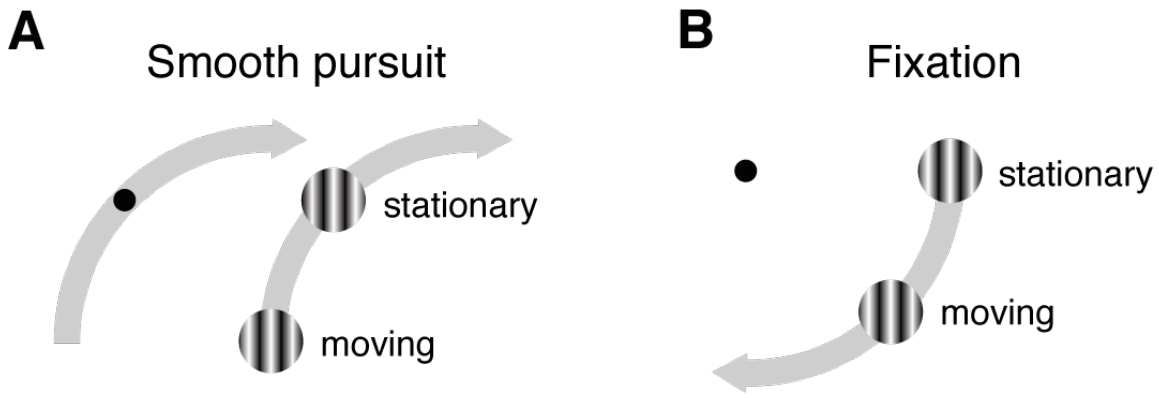


Figure 6. Experiment 2. Experimental layout. A) The smooth pursuit condition was identical to that of experiment 1. B) In the fixation condition the target (sinusoidal grating) remained either stationary on the horizontal meridian, to the right of the fixation spot (black dot), or moved on a circular trajectory (gray curved arrow) in the lower right quadrant of the visual field. In both conditions targets were defined according to their retinal image motion as either ‘moving’ or ‘stationary’.

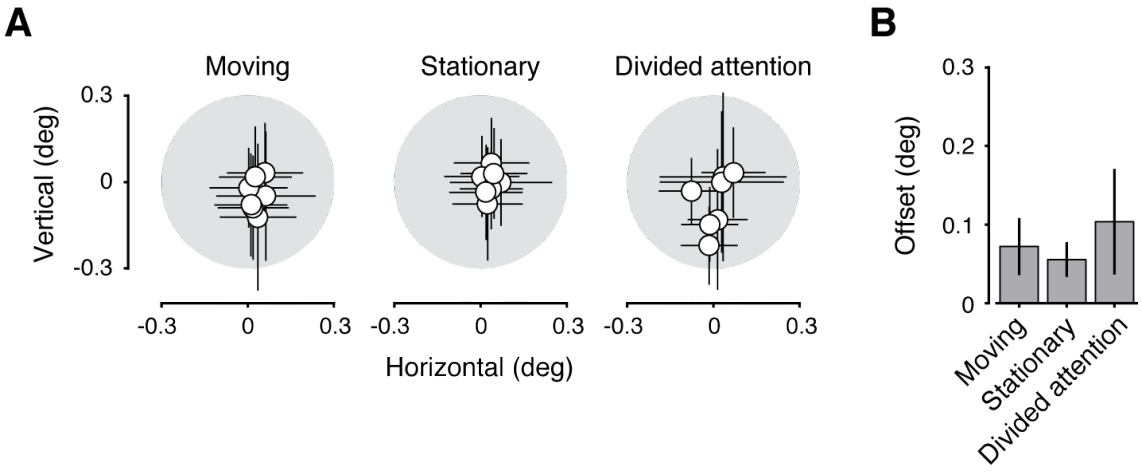


Figure 7, Experiment 2. Eye positions during fixation. A) Average eye positions (white circles) of individual subjects (*n* = 7) relative to the fixation spot (gray disc) in the three conditions. B) Average offset of fixation positions. The offset represents the distance from the fixation spot center to the mean average eye position of a subject. All data represent mean ± 1Std.

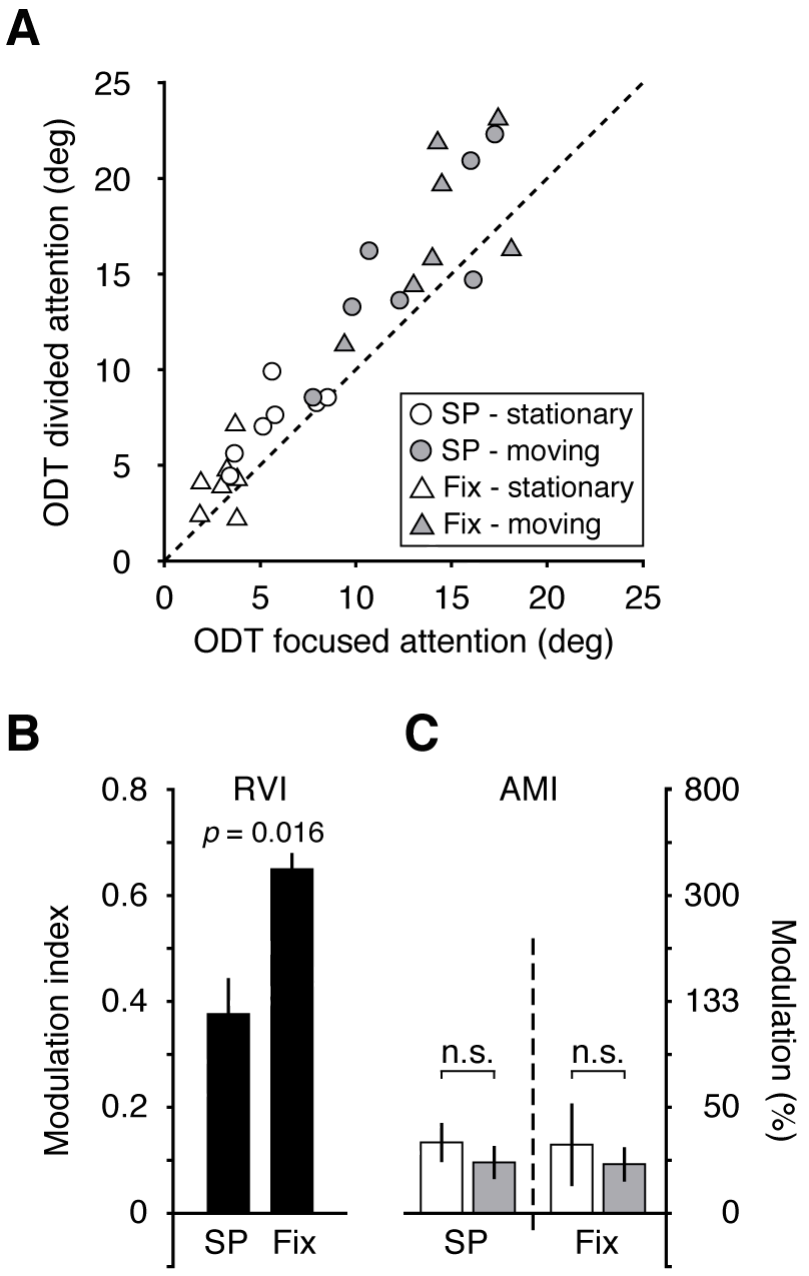


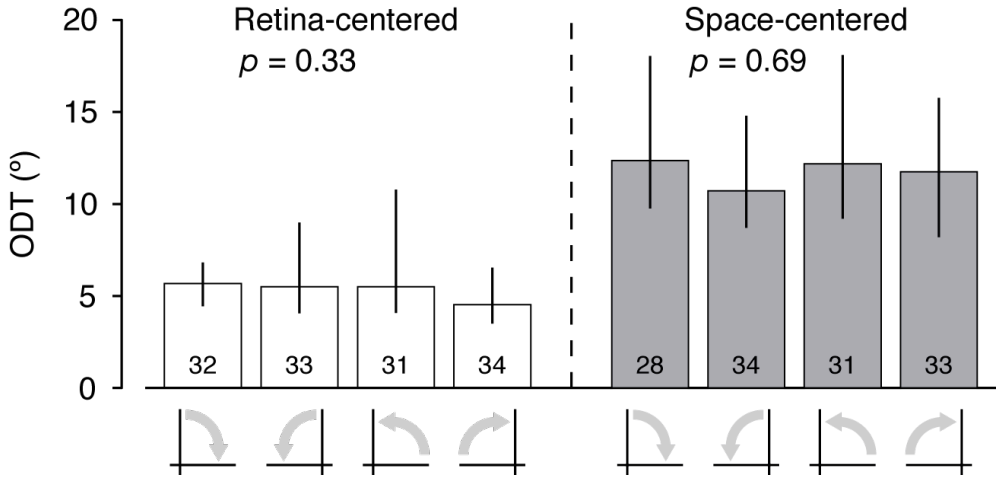
Figure 8. Experiment 2. Effect of dividing attention during fixation (Fix) and smooth pursuit (SP). A) Raw ODTs of individual subjects. Average ODTs with focused attention on the stationary (white) or the moving (gray) target are plotted against corresponding ODTs during divided attention for both smooth pursuit (circles) and fixation (triangles). B) Average RVIs. C) Average AMIs. Color-coding is similar to that in 'A'. All error bars represent SEM.

Auxiliary Document

Effect of smooth pursuit direction and screen quadrant on ODTs

We tested whether the direction of the smooth pursuit movement (from horizontal to vertical meridian or vice versa) and the quadrant of the screen across which the smooth pursuit target was moving (upper left or upper right) influenced the subjects' performance.

AF Figure 1 shows the median ODT values for the four possible combinations of pursuit direction and screen quadrant for both reference frames. Only the data of subjects that were tested in all possible combinations were included in the analysis (n = 9). A Kruskal-Wallis ANOVA, conducted separately on the data of each reference frame, showed that neither smooth pursuit direction nor the quadrant of the screen affected ODTs.



AF Figure 1. Experiment 1. Median ODT values. The bars represent individual combinations of smooth pursuit direction, screen quadrant and target reference frame (labels on the abscissa). Arrows indicate pursuit direction and angles pursuit quadrants. The number of included staircases (out of a maximum of 36) is indicated on each bar. Some staircases were discarded following the exclusion criteria. The error bars represent the range between the 25th and 75th percentile of the distributions. The p values correspond to Kruskal-Wallis ANOVAs across the four combinations for each reference frame.

Eye position calibration measurements

In order to quantify the measurement error of the smooth pursuit signals and the accuracy of the saccade detection algorithm we conducted calibration measurements in three of the subjects ('jcm', 'rni', 'thl').

METHODS

Procedure

The subjects were instructed to pursue the pursuit spot moving at the same speed and trajectory as in experiment 1 and 2 (AF Figure 2A). During the smooth pursuit period (930-1730 ms after smooth pursuit onset) we presented a small black square (*area* = 0.4 degrees visual angle ^{2}) at one of four possible distances from the pursuit spot (1, 2, 3 and 8.75 degrees visual angle). The targets presented at 1, 2 and 3 degrees visual angle moved at trajectories parallel to the smooth pursuit spot and the target presented at 8.75 degrees visual angle remained fixed at the screen center, resembling the space-centered target of experiments 1 and 2. Similar to experiment 2, smooth pursuits were restricted to the upper left quadrant of the screen and the direction was from horizontal to vertical meridian. During the smooth pursuit movement subjects made a saccade towards the saccade target as soon as they detected it and then a second saccade to return to the smooth pursuit target in order to finish the smooth pursuit movement. We also included trials during which no saccade target was presented and subjects simply performed the smooth pursuit ('no-saccade' trials). Each subject completed ten trials per condition. Trials of the different conditions were randomly intermixed within the experimental block.

Analysis of eye position data

The methods for recording and processing of the eye position data were identical to those described for the original experiments.

Results

Saccade velocity

The main point of this experiment was to define a velocity threshold for the detection of saccades that might have contaminated the smooth pursuit data.

AF Figure 2B shows the spatial traces of example trials of one subject ('rni') for the different saccade/no-saccade conditions. The eye position trace of the 'no-saccade' trial (light blue) superimposes the trajectory of the smooth pursuit dot (black), indicating that the subject

pursued the target with high spatial accuracy. The spatial traces of the saccade trials (dark blue, purple, orange, green) clearly show the amplitude of the saccades on top of the smooth pursuit. All saccade end points are located near the respective saccade target trajectory (grey), revealing the degree of spatial precision of the saccadic eye movement.

For each trial we computed the velocity as described in the article. AF Figure 2C illustrates the velocity profiles of the trials shown in AF Figure 2B. The velocity of the 'no saccade' trial (light blue, upper panel) oscillates around the smooth pursuit target velocity (dashed line). However, the fluctuations do not seem to show a clear bias to lower or higher velocities. These small fluctuations may indicate intrinsic noise in the eye movement measurements. Their maximal velocity was always lower than 20 degrees/sec (see AF Figure 2D). This may be taken as an indication of how noisy our eye movement measurements were and was taken into account for choosing a threshold criterion for our saccade detection method.

The velocity profiles of the saccade trials (lower panel) contain two peaks. The first one belongs to the saccadic eye movement towards the target and the second one to the saccade returning from the target to the smooth pursuit spot. We determined the peak velocity of individual trials and averaged those across trials with identical saccade distances (we did this also for 'no saccade' trials, see above). AF Figure 2D shows the typical relationship between saccade amplitude and peak velocity (Van Gisbergen et al. 1984) for each of the three subjects.

Since we were mainly interested in the detection of small saccades, we restricted the statistical analysis to the smallest saccade distance of 1 degree. In those trials the saccade target appeared close to the smooth pursuit spot (which had a diameter of 0.6 degrees).

The average peak velocities of the three subjects are remarkably similar (*jcm* = 28.25; *rni* = 28.93; *thl* = 29.8 degrees/s; $p = 0.68$, one-way ANOVA). Importantly, all of them were higher than 20 degrees/s ($p < 0.0001$, *t*-test), which was subsequently defined as the threshold criterion for our saccade detection algorithm.

Using the 20 degrees/s criterion, we could detect saccades in the saccade trials at a perfect score for all subjects. In the 'no-saccade' trials we detected none in two of the subjects and one (out of ten) in subject 'jcm'. We therefore conclude that our threshold criterion is appropriate for detecting saccades in our experiments.

Velocity during 'no-saccade' trials

In order to assess the measurement error of the smooth pursuit signals, we determined the subjects eye velocity during 'no-saccade' trials. Here, the subjects only pursued the dot.

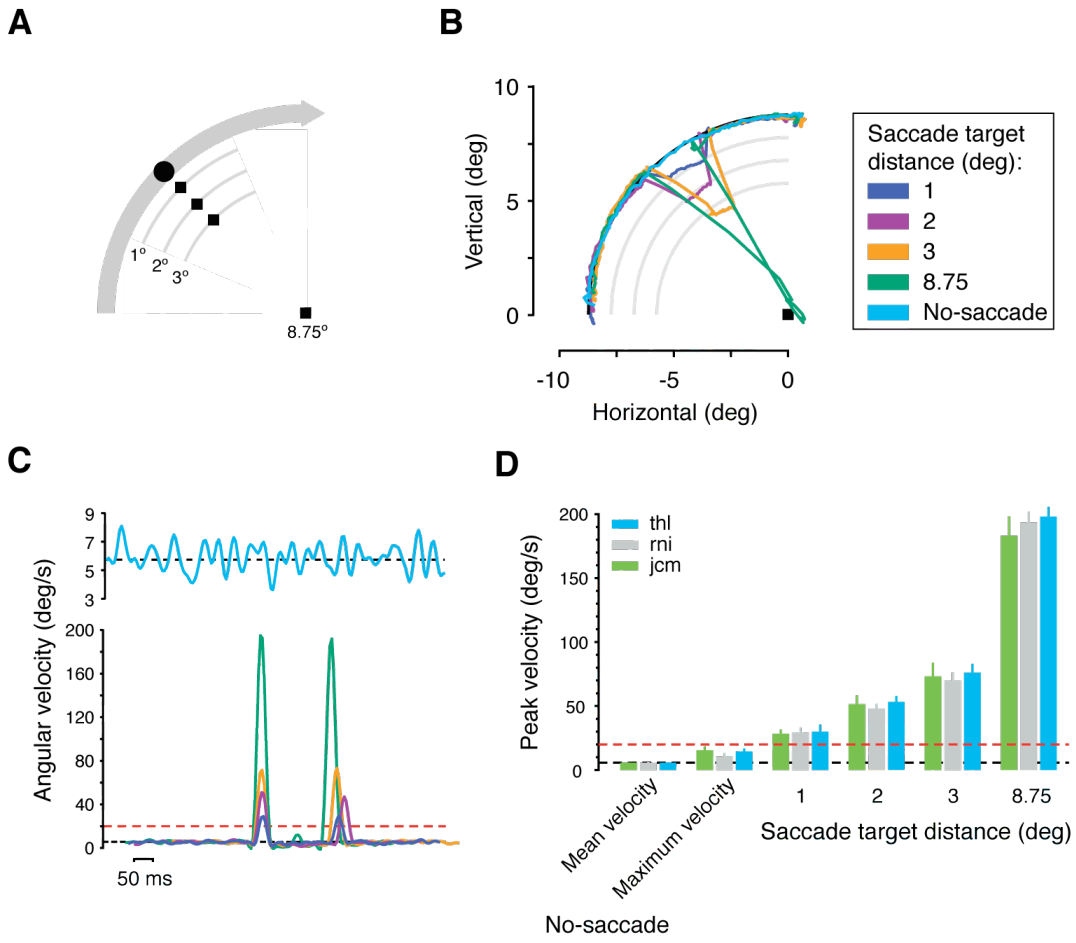
Therefore the standard deviation of the eye movement measurements provide a baseline measurement for comparison against trials where subjects performed a second task, i.e., covertly attended to the target grating(s). The average eye velocities during 'no saccade' trials are illustrated in AF Figure 2D.

We tested, for the three subjects, whether eye velocities during the no-saccade task were different from those during trials that required covertly attending to a grating, i.e. those in experiments 1 and 2. We computed an unpaired t-test for each subject using the gains of all available trials in the former two experiments and compared them to the gains of the 'no-saccade' trials. The following table summarizes the results:

AF Table 1. Comparison of smooth pursuit gains during trials of experiments 1 and 2 and 'no-saccade' trials of the present experiment. Data represent mean ±1Std.

Subject	Experiments 1 & 2 (n >100)	'No-saccade' (n = 10)	P unpaired t-test
jcm	1.023 ± 0.055	1.0 ± 0.028	0.18
rni	1.014 ± 0.06	0.989 ± 0.095	0.2
thl	1.032 ± 0.063	1.002 ± 0.049	0.13

These data suggests that covertly attending to the target grating or simply pursuing the target did not have a major influence on our measurements of smooth pursuit performance. We believe that the training provided to the subjects at the beginning of the sessions may have made the pursuit almost 'automatic' in our subjects (see also experiment 2 in the manuscript).



AF Figure 2. Saccade detection experiment. A) Task design. During smooth pursuit subjects made a saccade towards a target (squares) at one out of four possible locations at different distances from the smooth pursuit target (black disc). For three distances (1, 2 and 3 degrees) the target moved on a circular trajectory (grey curved lines) while for the largest distance (8.75 degrees) the target remained stationary at the screen center. In the 'no saccade' condition, only the smooth pursuit movement was required. B) Spatial eye position traces of example trials in the different conditions. The grey curved lines indicate the distances of the saccade target trajectories. C) Velocity profiles of the example trials in 'B' for the 'no-saccade' trial (upper panel) and the saccade trials (lower panel). The red dashed line indicates the velocity criterion (20 degrees/s) of the saccade detection algorithm and the black solid line the velocity of the smooth pursuit target. Color-coding follows conventions in 'B'. D) Average peak velocity. The error bars represent Std.

Effect of target retinal velocity on orientation discrimination performance

In this experiment, we tested whether the differences in ODTs found in experiment 1 for targets centered in the two reference frames could be caused by differences in the targets' retinal velocity. We designed an experiment that allowed us to change target velocity while keeping the eccentricity of a target grating constant and changing the target position in space and on the retina. We then measured the ODTs for different velocities of the target grating.

METHODS

Apparatus and Stimuli

Apparatus, viewing conditions, target grating and smooth pursuit dot were the same as in the previous experiment.

Subjects

One of the authors ('law') and two undergraduate students ('dim', 'jon') participated in the experiment.

Procedure

During trials, two stimuli, a pursuit dot and a sinusoidal target grating identical to those used in experiment 1, were presented. Upon trial initiation, the smooth pursuit dot appeared directly above the target and started moving horizontally across the screen at a velocity of 5.75 deg/s (AF Figure 3A). During trials, the target grating and the pursuit dot kept a constant distance of 8.75 degrees. Different target velocities were achieved by varying the orbiting (angular) displacement of the grating per time unit (AF Figure 3B). The observers' task was to pursue the dot while attending to the orbiting target and report, at the end of the trial, a change in its orientation (clockwise or counterclockwise, see previous experiment). Seven different target velocities were tested (in deg/s): 0; 1.0; 1.9; 2.9; 3.8; 4.8; 5.75. The task was run in blocks of 160 trials. Within a block, two different orbiting target velocities were presented (in deg/s): 0/1.9; 1.0/2.9; 2.9/4.8; 3.8/5.75. These velocities represent the target retinal velocity. The number of trials per target retinal velocity within a block was balanced and their occurrences randomized. Every subject completed the same block at least twice.

Measurements and analysis

We measured ODTs using the staircase method described in the previous experiment. Starting values for all staircases were between 3° and 9°. In each block, separate staircases were run for every combination of target velocity, starting value, and orientation change

direction (clockwise or counter-clockwise). In order to calculate the average ODT for a specific target velocity, we pooled data from similar staircases. Staircases with insufficient reversal points (< 5) were discarded ('law': 6.2%; 'jon': 7.8%; 'dim': 1.6%).

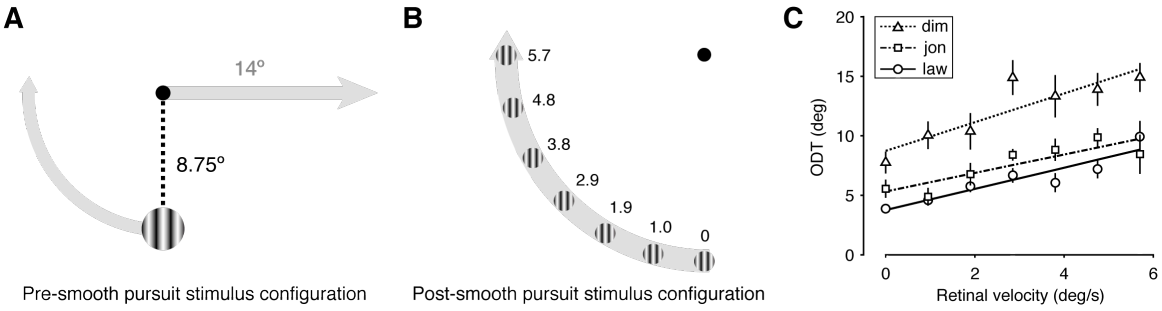
RESULTS

AF Figure 3C shows the three subject's ODTs as a function of target retinal velocity. We fitted regression lines to these data. The lines have significant positive non-zero slopes ('dim': *slope ± 95% confidence interval (CI)* = 1.2 ± 0.55, R^2=0.79; 'jon': *slope ± CI* = 0.78 ± 0.39, R^2=0.76; 'law': *slope ± CI* = 0.89 ± 0.32, R^2=0.86). In all cases the 95% confidence interval does not include zero, demonstrating that ODTs increased linearly as a function of the target retinal velocity.

Interestingly, for one of the subjects ('law') who also participated in experiment 1, the ODTs for the lowest retinal velocity (0 deg/s) was 5°, while for the highest retinal velocity (5.75 deg/s) it was 10°. These values are close to those obtained in experiment 1 for the retina centered target (5°), which also had a retinal velocity of 0 deg/s, and for the space centered target (14°), which had a retinal velocity of 5.75 deg/s. The small increase in ODTs in experiment 1 relative to those predicted by our measurements in this experiment may be due to at least two factors. First, the smooth pursuit trajectories in experiment 1 were circular, introducing an extra-level of difficulty relative to the present experiment where pursuit trajectories were straight. Second, in experiment 1 two target gratings were presented, while here only one grating was displayed, making the latter display less crowded and the task easier.

A factor that may explain this increase in ODTs is motion blur (Chung et al., 1996; Kelly, 1979; Land, 1999), which causes retinal images to spread over a spatiotemporal area thereby reducing the visibility of the stimulus' features. An orientation change in a stimulus moving on the retina might therefore be harder to perceive. Blurring can occur at velocities as low as 1 deg/s for high spatial frequency images (Land, 1999), matching our observation of ODT elevation at velocities of 0.95 deg/s. For lower retinal speeds such as those used by Khurana & Kowler (1987) — 0.42 deg/s or 0.83 deg/s, only comparable to the lowest velocity used in our study — this may not apply. Another factor that may contribute to our results is the use of low-level stimuli (sinusoidal gratings), which are likely represented in retinotopic brain areas with comparably small receptive fields, such as V4 (Gardner et al., 2008). Letters (used by Khurana & Kowler, 1987) on the other hand, are likely represented in higher-level brain areas, such as left occipito-temporal cortex (Grainger, Rey, & Dufau, 2008), or the posterior

parietal cortex (Cattaneo, Rota, Vecchi, & Silvanto, 2008), where neurons show position invariance for stimuli inside their receptive field (Duhamel et al., 1997; Galletti et al., 1993). In the latter case visual representations might be less susceptible to changes in retinal image velocity, at least within the spatial limits of their position invariance.



AF Figure 3. A) Experimental layout. Observers pursued the black dot while attending to the orbiting grating. See text for details. The dotted line shows the constant distance between the pursuit dot and the target grating. The straight grey arrow represents the trajectory of the smooth pursuit dot, and the circular arrow the orbiting trajectory of the grating. B) Different final positions of the target grating depending on its orbiting velocity (numbers represent degrees/s). Each position is depicted relative to the smooth pursuit dot (upper right corner). The grey arrow depicts the target trajectory on the retina. C) ODTs as a function of target retinal velocity for the three subjects. Data represent mean ODTs ±1SEM. The lines illustrate the best linear regression models fitted through the data.

References

Cattaneo, Z., Rota, F., Vecchi, T., & Silvanto, J. (2008). Using state-dependency of transcranial magnetic stimulation (TMS) to investigate letter selectivity in the left posterior parietal cortex: a comparison of TMS-priming and TMS-adaptation paradigms. *European Journal of Neuroscience, 28,* 1924-1929.

Chung, S. T., Levi, D. M., & Bedell, H. E. (1996). Vernier in motion: what accounts for the threshold elevation? *Vision Research, 36*, 2395-2410.

Duhamel, J. R., Bremmer, F., BenHamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature, 389*, 845-848.

Galletti, C., Battaglini, P. P., & Fattori, P. (1993). Parietal neurons encoding spatial locations in craniotopic coordinates. *Experimental Brain Research, 96*, 221-229.

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *Journal of Neuroscience, 28*, 3988-3999.

Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: from pixels to pandemonium. *Trends in Cognitive Sciences, 12*, 381-387.

Kelly, D. H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America, 69*, 1340-1349.

Khurana, B., & Kowler, E. (1987). Shared attentional control of smooth eye movement and perception. *Vision Research, 27*, 1603-1618.

Land, M. F. (1999). Motion and vision: why animals move their eyes. *Journal of Comparative Physiology [A], 185*, 341-352.

Van Gisbergen, J. A. M., Van Opstal, J. A. & Ottes, F. P. (1984). Parametrization of saccadic velocity profiles in man. In: Gale AG, Johnson F (eds) Theoretical and applied aspects of eye movement research. *Elsevier Science*, New York, 87–94.

2.3 Attention Differentially Modulates Similar Neuronal Responses Evoked by Varying Contrast and Direction Stimuli in Area MT

Two main categories of models have been proposed to account for the effects of attention on the activity of visual neurons. First, the 'response gain' model, which proposes a scaling or additive modulation of a neuron's firing rate independently of the attended stimulus features and contrast. Here, the firing rate will depend on the degree of attentional effort, and the allocation of attention to features or spatial locations. Second, the 'contrast gain' model, which proposes that the effects of attention are also dependent on the contrast of the stimulus, being larger for stimuli with intermediate relative to low and high contrast. In the present study, we distinguished between these two alternatives by recording the effects of attention on neuronal responses in area MT to stimuli, changing in contrast and direction while keeping the allocation of spatial and feature-based attention constant.

Two moving RDPs were positioned inside the neurons receptive field. One of them always had high contrast and moved in the neurons' anti-preferred direction (AP-pattern). The second, could have two different configurations across trials: a) it moved in the neuron's preferred direction but had from trial to trial different contrast levels, or b) it had the same contrast as the AP-pattern but moved from trial to trial in different directions. Either the contrast or the motion direction of the test pattern was adjusted in such a manner that various configurations of AP-pattern and test pattern evoked similar response levels when the animal ignored both RDPs (fixation or sensory response).

In such a situation, the 'response gain' model would predict that attending to the AP-pattern results in a response suppression that is similar for both configurations, i.e., changing contrast and changing directions of the test pattern. On the other hand, the 'contrast gain' model would predict a stronger suppression for test patterns with intermediate relative to low and high contrast.

We found that when the monkeys directed attention to the AP-pattern, the neurons' response was suppressed, and that indeed the magnitude of this suppressive effect changed depending on the contrast of the unattended test pattern (higher for intermediate relative to high and low contrast test patterns). Additionally, we found that the motion direction of the test pattern modulated the intensity of the attentional modulation, an effect attributable to feature-based attention, modulating the strength of input signals into MT neurons.

These results are incompatible with a scaling or additive modulation of MT neurons' response by attention, as suggested by 'response gain' models. On the other hand, they agree with recently proposed contrast gain models of attention, in which space-, and feature-based attention modulate the strength of input signals into a neuron's normalization circuit, and therefore the strength of inhibitory mechanisms.

Behavioral/Systems/Cognitive

Attention Differentially Modulates Similar Neuronal Responses Evoked by Varying Contrast and Direction Stimuli in Area MT

Paul S. Khayat, Robert Niebergall, and Julio C. Martinez-Trujillo
Cognitive Neurophysiology Laboratory, Department of Physiology, McGill University, Montreal, Quebec H3G 1Y6, Canada

The effects of attention on the responses of visual neurons have been described as a scaling or additive modulation independent of stimulus features and contrast, or as a contrast-dependent modulation. We explored these alternatives by recording neuronal responses in macaque area MT to moving stimuli that evoked similar firing rates but varied in contrast and direction. We presented two identical pairs of stimuli, one inside the neurons' receptive field and the other outside, in the opposite hemifield. One stimulus of each pair always had high contrast and moved in the recorded cell's antipreferred direction (AP pattern), while the other (test pattern) could either move in the cell's preferred direction and vary in contrast, or have the same contrast as the AP pattern and vary in direction. For different stimulus pairs evoking similar responses, switching attention between the two AP patterns, or directing attention from a fixation spot to the AP pattern inside or outside the receptive field, produced a stronger suppression of responses to varying contrast pairs, reaching a maximum (~20%) at intermediate contrast. For invariable contrast pairs, switching attention from the fixation spot to the AP pattern produced a modulation that ranged from 10% suppression when the test pattern moved in the cells preferred direction to 14% enhancement when it moved in a direction 90° away from that direction. Our results are incompatible with a scaling or additive modulation of MT neurons' response by attention, but support models where spatial and feature-based attention modulate input signals into the area normalization circuit.

Introduction

It has been established that directing attention to the spatial location, or to nonspatial features of visual stimuli, modulates the responses of visual cortical neurons of primates (Desimone and Duncan, 1995; Treue, 2001; Reynolds and Chelazzi, 2004; Maunsell and Treue, 2006). Some studies have described the modulation caused by spatial attention as a scaling or upward shift of a neuron's response function for attributes such as motion direction (Treue and Martínez Trujillo, 1999), orientation (McAdams and Maunsell, 1999), and contrast (Williford and Maunsell, 2006; Thiele et al., 2009). Conversely, other studies have reported that spatial attention modulates more strongly responses to intermediate-contrast stimuli, producing a shift in the response function along the contrast axis (Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002). The first type of modulation can be described as stimulus/contrast independent, and the second as contrast dependent.

A stimulus/contrast-independent modulation may suggest a mechanism that changes the neuron's response independently of the type (e.g., motion direction, orientation) or strength (luminance/contrast) of the sensory input. Alternatively, a contrast-dependent modulation may suggest a process wherein attentional effects are constrained by contrast gain control mechanisms, e.g., attention may modulate the strength of inputs into a neuron, producing an effect similar to changes in stimulus contrast (Reynolds and Heeger, 2009). Some studies have attempted to discriminate between these alternatives by comparing the goodness of fit of model equations to single-cell responses evoked by varying contrast stimuli (Martínez-Trujillo and Treue, 2002; Williford and Maunsell, 2006). They have reported marginal differences between the goodness of fit of the models, with several models explaining most of the variance in the data. Moreover, factors such as differences in the studies' experimental design may have contributed to the reported heterogeneity in results.

In humans, some fMRI studies have described contrast-independent attentional modulation of BOLD signals in visual cortex (Buracas and Boynton, 2007; Murray, 2008), while others have reported contrast-dependent modulations (Li et al., 2008; see also Ekstrom et al., 2009 in the monkey). Behavioral studies have also described both modulation types of psychometric functions (Carrasco, 2006). Thus, the heterogeneity in results also applies to human studies. Recent modeling work has aimed at reconciling the various findings (Ghose, 2009; Lee and Maunsell, 2009; Reynolds and Heeger, 2009). However, experimental data directly clarifying the issue are scarce. Moreover, it remains unknown whether different forms of attention (e.g., spatial and feature based) produce different effects on neuronal responses to varying contrast stimuli.

Received Oct. 27, 2009; revised Dec. 14, 2009; accepted Dec. 25, 2009.

This study was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) fellowship awarded to P.S.K. and Canada Foundation for Innovation, Canadian Institutes of Health Research, and NSERC grants awarded to J.C.M.-T. We thank Dr. E. P. Cook and Dr. C. Pack for providing valuable comments on the manuscript. We thank W. Kucharski and S. Nuara for technical assistance.

Correspondence should be addressed to Julio C. Martinez-Trujillo, Department of Physiology, McGill University, Room 1220, 3655 Prom. Sir. W. Osler, Montreal, QC H3G 1Y6, Canada. E-mail: julio.martinez@mcgill.ca.

DOI:10.1523/JNEUROSCI.5314-09.2010

 0270-6474/10/302188-10$15.00/0

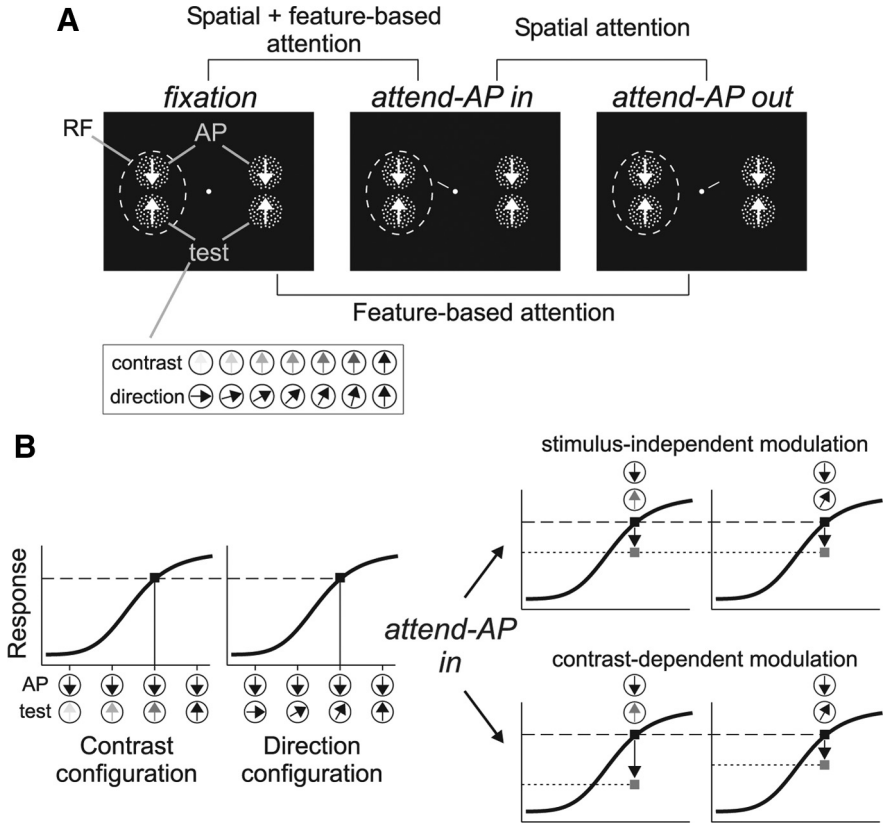


Figure 1. Task design. ***A***, Stimulus display in the three behavioral conditions and the different comparisons. Each panel shows two pairs of high-contrast RDPs presented inside and outside the RF (dashed circle), and moving in the neuron's antipreferred (AP pattern) and preferred (test pattern) direction. Insert at the bottom illustrate the different contrasts and motion directions of the test pattern. ***B***, The graphs on the left illustrate hypothetical response profiles of an MT direction-selective neuron as a function of changes in contrast or motion direction of the test pattern (illustrated on the abscissa). Black squares mark similar response levels evoked by different stimuli. The graphs on the right illustrate the predictions of the stimulus-independent and contrast-dependent modulation when attention is shifted from the AP pattern outside the RF to a similar pattern inside (arrows and gray squares) (see Materials and Methods for details).

Here, we investigated these issues by comparing the modulation produced by spatial and feature-based attention on the responses of macaque MT neurons to stimuli that varied in direction and contrast but evoked similar responses when unattended. We reasoned that when animals attend to the same location and feature, a stimulus/contrast-independent modulation would produce similar response changes. On the other hand, a contrast-dependent modulation would produce larger changes in responses evoked by stimuli with intermediate contrast. Our data favor the latter explanation, and further show that the modulation produced by spatial, feature-based attention, and their interaction share the same constraints.

Materials and Methods

Two male macaque monkeys participated in the experiments. Monkey Lu weighed 5.5 kg and monkey Se 7.2 kg. Standard electrophysiological procedures were used to record neuronal activity in area MT. All surgical procedures for the implant of recording chambers and head holders were conducted under general anesthesia (see Martinez-Trujillo and Treue, 2004). The health of the animals was carefully and periodically monitored by an experienced veterinarian, and the animals were provided with behavioral enrichment protocols when returning to their home cages. All procedures complied with the Canadian Council of Animal Care guidelines and were approved by the McGill animal care committee.

Behavioral task. The animal was required to keep fixation within a 1.5° window centered on a small spot (0.1 degree^2), and then initiate a trial by pressing a button. After 470 ms, two pairs of random dot patterns (RDPs) appeared, one located inside the receptive field (RF) of the recorded MT neuron, and the other located outside, in the opposite hemifield (Fig. 1*A*). Each pair consisted of a high-contrast RDP moving in the neuron's antipreferred direction (AP pattern) and a second RDP (test pattern) that (1) moved in the neuron's preferred direction but could have, from trial-to-trial, different contrast levels (contrast configuration), and (2) had the same contrast as the AP pattern but could move, from trial-to-trial, in different directions (direction configuration) (Fig. 1*A*, inset).

We tested three different task conditions. In the "fixation" condition, the animal had to detect a subtle luminance change in the fixation spot, which occurred at a random time between 1010 and 3250 ms after stimulus onset. In the two other conditions, a small line (1° length) appeared next to the fixation spot 350 ms after stimulus onset. This cue-line pointed toward one of the AP patterns, thereby instructing the monkey to direct attention either inside ("attend-AP in" condition) (Fig. 1*A*, middle) or outside ("attend-AP out" condition) (Fig. 1*A*, right) the RF. The target underwent a brief direction change 660–2900 ms after cue onset (23° during 100 ms). The animal had to release the button within a response time window of 150–500 ms after the change to receive a juice reward. The direction change intensity was chosen in such a way that the animals' performance in most sessions was above 75% of correct detections.

By comparing responses in the different conditions, we could isolate the effects of spatial and feature-based attention and their combination (Fig. 1*A*). When the animal switches attention between the AP pattern outside the RF to the other identical AP pattern inside, only the allocation of spatial attention changes since the attended feature remains the same. When the animal switches attention between the fixation spot and the AP pattern outside the RF, spatial attention remains outside the RF. However, in the feature domain, it is directed from a condition where no motion feature is attended ("fixation") to a condition where the neuron's antipreferred direction is attended. This manipulation has been previously used to isolate feature-based attentional effects in MT neurons (Martinez-Trujillo and Treue, 2004). Finally, switching attention from the fixation spot to the AP pattern inside the RF combines the effect of directing attention into the RF (spatial attention) with the effect of directing attention to the antipreferred direction (feature-based attention).

To guarantee that during the attended conditions the monkey focused attention on the target, the uncued AP pattern located in the opposite visual hemifield changed direction in half of the trials. The monkey had to ignore this distractor change and wait until the target changed. Trials in which the monkey responded to the distractor change or broke fixation were terminated without reward. The different trial types were presented in random sequence, and both animals performed between 6 and 15 trials per stimulus type in each behavioral condition.

Stimuli. The stimuli were back-projected on a screen by a video projector (NEC WT610, 1024 × 768 pixels resolution at 85 Hz). We used a viewing distance of 57 cm. The RDPs were generated by plotting bright dots (dot size = 0.01 degree^2) at a density of 4 dots per degree^2 within a circular stationary virtual aperture on a dark background (luminance = 0.02 cd/m^2). Dots moved with 100% coherence at the preferred speed of the neurons. When they crossed the aperture's border they were replot-

ted at the opposite side. The size of the RDP (1.3–3° diameter) was chosen so that two RDPs fitted inside the boundaries of the classical RF excitatory region.

Stimulus contrast was quantified as the SD of local luminance values (Moulden et al., 1990), and computed as previously described for similar configurations of RDPs (Martínez-Trujillo and Treue, 2002). The values were expressed in percentage of the highest contrast value (i.e., AP-pattern contrast = 100%). In the contrast configuration, we used different contrast levels of the test pattern obtained by manipulating the luminance of this pattern's dots while keeping constant the background and the AP-pattern luminance. The contrast levels of the test pattern relative to the AP pattern were 0.02, 0.1, 0.3, 0.7, 1.5, 14, and 100%. In the direction configuration, we used seven different motion directions of the test pattern (departing from the recorded neuron's preferred direction in steps of 15° until 90° away). In this configuration, the contrast of the test pattern was always 100% (Fig. 1*A*).

Recordings and data analysis. Transdural penetrations were made with guide tubes through a chamber implanted on top of a craniotomy of the parietal bone and providing access to area MT. Spikes were recorded using extracellular tungsten electrodes (1–2 MΩ at 1 kHz, FHC), and a Plexon data acquisition system (Plexon). The electrode signal was amplified and filtered before being digitized at 40 kHz. Single unit activity was isolated using a window discriminator. An interactive stimulus presentation program was used to qualitatively assess the location and size of the neurons' RF, as well as their preferred direction and speed. Cells were determined to be MT units according to their response properties (i.e., direction selectivity and RF position and size), as well as by the position of the electrode relative to the superior temporal sulcus, as localized through MRI scans (Fig. 2). During the recordings, an infrared eye-tracking device (EyeLink) was used to monitor eye position at a sampling frequency of 200 Hz.

The neuron's response was determined by averaging the firing rate across trials within a 600 ms period that started 150 ms after cue onset and did not include the stimulus change. We measured neuronal activity during the sustained response period to isolate the modulation produced by endogenous (voluntary) attention, i.e., after the cue was provided to the animals and they directed attention to the target stimulus (Khayat et al., 2006; Busse et al., 2008). This avoids including in our analysis period potential effects of exogenous attention produced by the stimulus/motion onset transient (Treue and Martínez Trujillo, 1999), as well as underestimates of the firing rate during the early phase of the trial produced by the variability in response latency associated with changes in the test stimulus contrast (Reynolds and Desimone, 2003; Lee et al., 2007). Cells were included in the analyses if at least six correctly performed trials per stimulus and condition were available (median = 12 trials; first and third quartiles equal 8 and 14 trials, respectively). A total of 101 units (65 in monkey Se and 36 in monkey Lu) were included in the analyses, and were recorded during 96 sessions (in 5 sessions two units were simultaneously isolated from the same electrode).

Population responses were obtained by first, normalizing each cell response to the response evoked by the optimal stimulus during the "fixation" condition, and then, pooling across neurons. The optimal stimulus was defined as the combination of the AP pattern and a full (100%) contrast test pattern moving in the cell's preferred direction. To quantify the strength of attentional effects between responses in the different conditions, we computed a modulation index (MI) using the following equation: $(R_{c1} - R_{c2})/(R_{c1} + R_{c2})$, where R_{c1} and R_{c2} are the responses in the two conditions to be compared (Treue and Martínez Trujillo, 1999). To extract, for each neuron, equated (similar) responses between stimuli in the direction and contrast configuration, the response evoked by each stimulus pair (in a 600 ms analysis period) was first normalized to the response evoked by the optimal stimulus pair, which was common to both configurations. We then equated the responses in the contrast configuration during either the "attend-AP out" or "fixation" condition to those in the same condition of the direction configuration, i.e., by finding corresponding response levels between the two configurations. This allowed obtaining equated normalized responses and computing the corresponding MIs for each response level. ANOVA and *t* tests were used to test for significant effects in the MIs distribution.

We also determined the modulation in task performance by computing a performance modulation index, perfMI = $(Perf_{c1} - Perf_{c2})/(Perf_{c1} + Perf_{c2})$, where $Perf_{c1}$ and $Perf_{c2}$ are the hit rates (proportion of correct direction change detections) corresponding to the two conditions to be compared. Ninety-five percent confidence intervals for each perfMI were computed using a bootstrap procedure (10,000 resampling). Comparisons between perfMIs were conducted using nonparametric Kruskal–Wallis ANOVA and sign tests.

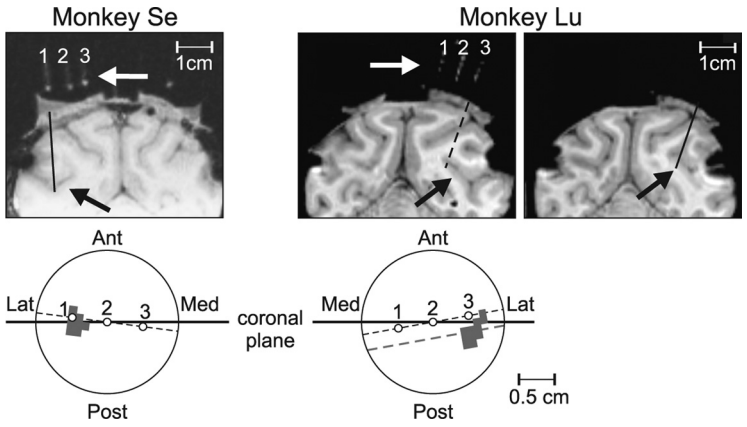


Figure 2. Recording location. MRIs from two monkeys showing the location of the recording chamber giving access to the superior temporal sulcus with putative MT recording sites (black arrows). The white arrows show glass capillaries filled with oil placed inside the chamber signaling the orientation of the electrode penetrations relative to the brain surface. The lower panels illustrate a top view of each monkey's chamber with the location of the capillaries (open circles) and the penetration sites (gray area). Ant, Anterior; Post, posterior; Lat, lateral; Med, medial. Note that the plane of the MRIs showing the capillaries (dashed black line in the lower panel) is slightly rotated with respect to the coronal plane (solid horizontal line crossing the center of the chambers). In monkey Se, the penetrations were made within the region surrounding capillary #1 (see MRI, and gray shaded area in the top view of the chamber). In monkey Lu, the penetrations were made posterior to capillary #3. The left MRI illustrates a slice through a plane where the three capillaries were aligned. The right MRI illustrates a slice through the plane—parallel to the capillaries' plane— crossing the center of the penetration area (dashed gray line and shaded region in the chamber's sketch). The solid black line in the MRI indicates a penetration reaching area MT. Black spots around the surface of the cortex are artifacts due to the titanium screws.

Results

Attentional modulation as a function of response level

The main idea behind our experimental design is illustrated in Figure 1*B*. The left panels display hypothetical response profiles (i.e., with attention directed outside the neuron's RF) of an MT direction-selective unit to two different stimulus configurations resembling the ones used in our experiment, each consisting of two moving RDPs presented inside the cell's RF. The changes in the test-pattern contrast or direction produce similar response profiles. However, resulting from different stimuli. For example, the AP pattern paired with a low-contrast test pattern moving in the cell's preferred direction will evoke the same response as when paired with a high-contrast test pattern moving in a direction different from preferred.

This experimental situation leads to clear predictions on the effects of attention on neuronal responses. If attention produces a response modulation independent of the stimulus type (direction

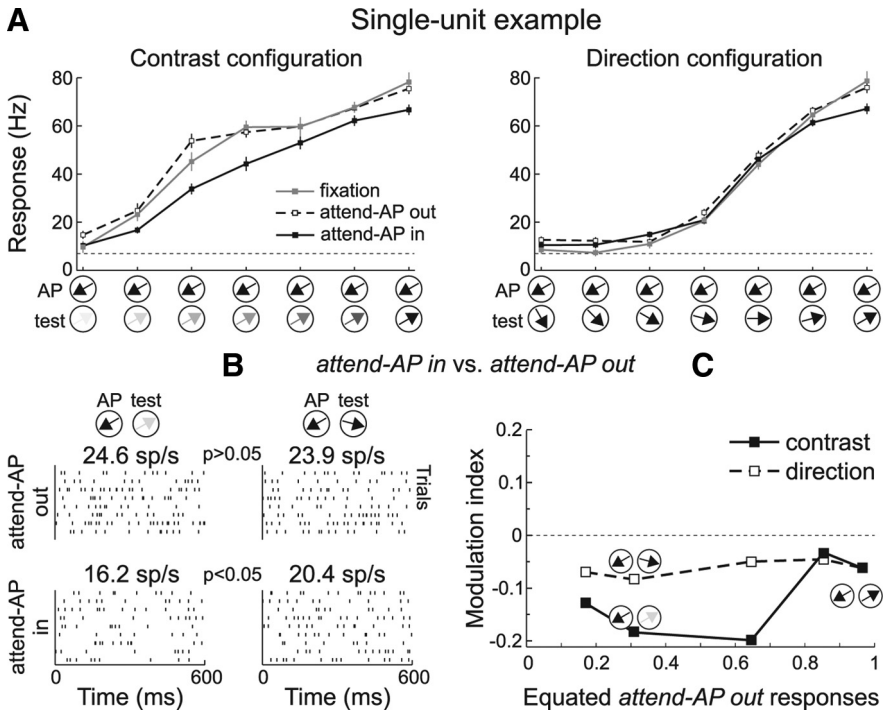


Figure 3. Attentional modulation of equated responses: single-unit example. ***A***, Average responses to different combinations of AP and test-pattern contrasts (left) and directions (right) in the "fixation" (gray), "attend-AP in" (black), and "attend-AP out" (dashed) conditions. The stimuli are illustrated along the abscissa. The dashed horizontal line indicates spontaneous response without stimulus. The neuron's preferred direction was 60° clockwise from vertical. Error bars, SEM. ***B***, Raster plots of the neuron's activity, and mean firing rates during trials with different stimulus pairs, in the "attend-AP out" (top) and "attend-AP in" (bottom) conditions. The *p* value indicates the significance level (*U* test) of differences between responses evoked by different stimuli. ***C***, MI across responses equated during the "attend-AP out" trials. Black squares and solid line, MI for responses to stimuli of the contrast configuration. Open squares and dashed line, MI for responses to stimuli of the direction configuration.

or contrast) but dependent on the response level, when switching attention between the two AP patterns the neuron's response to both stimulus combinations should be modulated by a similar amount (Fig. 1*B*, top right). As an alternative, a contrast-dependent modulation predicts that the response change will be larger for the combination with the lower contrast (Fig. 1*B*, bottom right). Note that this design keeps the allocation of attention constant during both, the contrast and the direction configuration trials (at the same location and on the same feature), regardless of changes in the test pattern identity, and without substantially altering task difficulty. It also groups features of two previous studies, one reporting multiplicative changes in response with attention (Treue and Martínez Trujillo, 1999), and the other reporting a contrast-dependent attentional modulation (Martínez-Trujillo and Treue, 2002).

We recorded the responses of 101 MT single neurons to the two stimulus configurations in two macaques during task trials of the three different conditions, "fixation," "attend-AP in," and "attend-AP out" (see Materials and Methods). Figure 3*A* shows responses of a single-unit to different stimulus pairs of the contrast (left) and direction configuration (right). Responses during "fixation" trials (gray line) were strongest when the test pattern had high contrast and moved in the preferred direction (rightmost data point in each panel), and became weaker as the test pattern decreased contrast, or changed direction away from preferred. Visual inspection of these data suggests that the magnitude of the differences in response between the "attend-AP in" (black) and "attend-AP out" (dashed line) conditions varied between the two stimulus configurations.

We investigated whether directing spatial attention inside the RF produced a similar effect on the same response level evoked by distinct stimulus combinations. Figure 3*B* shows raster plots of the neuron's response to two different stimulus pairs. Responses during the "attend-AP out" condition were similar when the test pattern had intermediate contrast and moved in the preferred direction (contrast pair, top left raster), and when it had high contrast and moved in a direction 45° away from preferred (direction pair, top right raster) (response = 24.6 ± 3.5 SEM and 23.9 ± 2.4 SEM spikes/s for the contrast and direction pairs, respectively; $p > 0.05$, *U* test). When directing attention to the AP pattern inside the RF, the stimulus-independent modulation predicts that responses to both pairs will be modulated in the same manner. However, we found that in the "attend-AP in" condition (bottom rasters) the response to the contrast pair was significantly suppressed (16.2 ± 1.4 SEM spikes/s, $p < 0.05$, *U* test), while the response to the direction pair did not significantly change (20.4 ± 1.3 SEM spikes/s, $p > 0.05$, *U* test).

In the same unit, we normalized responses evoked by all stimulus pairs to the response evoked by the combination of AP pattern and maximal contrast preferred test pattern during "fixation" (black rightmost data point in both panels of Fig. 3*A*). We then identified similar response levels during trials of the "attend-AP out" condition that were evoked by different stimulus combinations (equated normalized responses between contrast and direction combinations, see Materials and Methods). To determine the strength of the modulation for each response level, we computed a modulation index $MI = (R_{in} - R_{out})/(R_{in} + R_{out})$, where R_{in} and R_{out} are the responses to each stimulus combination in the "attend-AP in" and "attend-AP out" conditions, respectively (Fig. 3*C*). Across different levels of equated "attend-AP out" responses, attending to the AP stimulus inside the RF suppressed more strongly responses to contrast (black squares) compared to direction (open squares) combinations.

We conducted the same analyses for each recorded neuron ($n = 101$), and then averaged the MIs across units within each equated "attend-AP out" normalized response level (Fig. 4). Overall, spatial attention (Fig. 4*A*) ("attend-AP in" vs "attend-AP out") produced a stronger modulation in the contrast configuration. The MIs were more negative for responses evoked by contrast (black squares and solid line) relative to direction (open squares and dashed line) stimulus pairs. The stars on the abscissa in each panel indicate the equated response levels with significant differences in MIs between trials of both configurations ($p < 0.05$, paired *t* test).

We further investigated whether feature-based attention and its combination with spatial attention produced a similar modulation. We equated responses during "fixation," and compared them to those of the "attend-AP in" and "attend-AP out" conditions. Note that the "fixation" condition is considered in this context as feature-neutral, since attention is away from a mo-

tion feature and from the RF (Fig. 1*A*) (Martinez-Trujillo and Treue, 2004). We found that the MIs for the comparison "attend-AP in" versus "fixation" (spatial + feature-based attention) (Fig. 4*C*) [MIs = $(R_{in} - R_{fix})/(R_{in} + R_{fix})$], and "attend-AP out" versus "fixation" (feature-based attention) (Fig. 4*E*) [MIs = $(R_{out} - R_{fix})/(R_{out} + R_{fix})$] were significantly more negative for stimuli of the contrast configuration ($p < 0.05$, paired *t* test) (Fig. 4*C*,*E*, stars).

Overall, these data show that at the level of the population, responses evoked by contrast pairs were more strongly suppressed by spatial attention ("attend-AP in" vs "attend-AP out"), feature-based attention ("attend-AP out" vs "fixation"), and the interaction of both types of attention ("attend-AP in" vs "fixation") than similar responses evoked by direction pairs. The differences in the modulation strength were in some cases as large as 20% (third and fourth data points from right to left in Fig. 4*C*).

One alternative explanation for the effects isolated in Figure 4, *A*, *C*, and *E*, is that they were caused by variations in the animals' attentional effort, due to differences in task difficulty. We consider this possibility very unlikely since in general, increases in attentional effort increase the attentional modulation (Spitzer et al., 1988; Ghose and Maunsell, 2002; Boudreau et al., 2006). In our scenario, it only makes sense to hypothesize that low-contrast test patterns were "easier to filter out" relative to the high-contrast direction stimuli leading to less attentional effort. Our results show the largest modulation for the lower contrast stimuli, which is exactly the opposite as predicted by this hypothesis. Nevertheless, we investigated this possibility by computing the animals' performance for trials that contributed to each equated response level, and expressed it as perfMI (see Materials and Methods). Changes in attentional effort should be reflected in the animals' performance, and if they caused the isolated differences in response modulation between conditions and stimulus configurations, the perfMIs should be correlated with the corresponding attentional MIs.

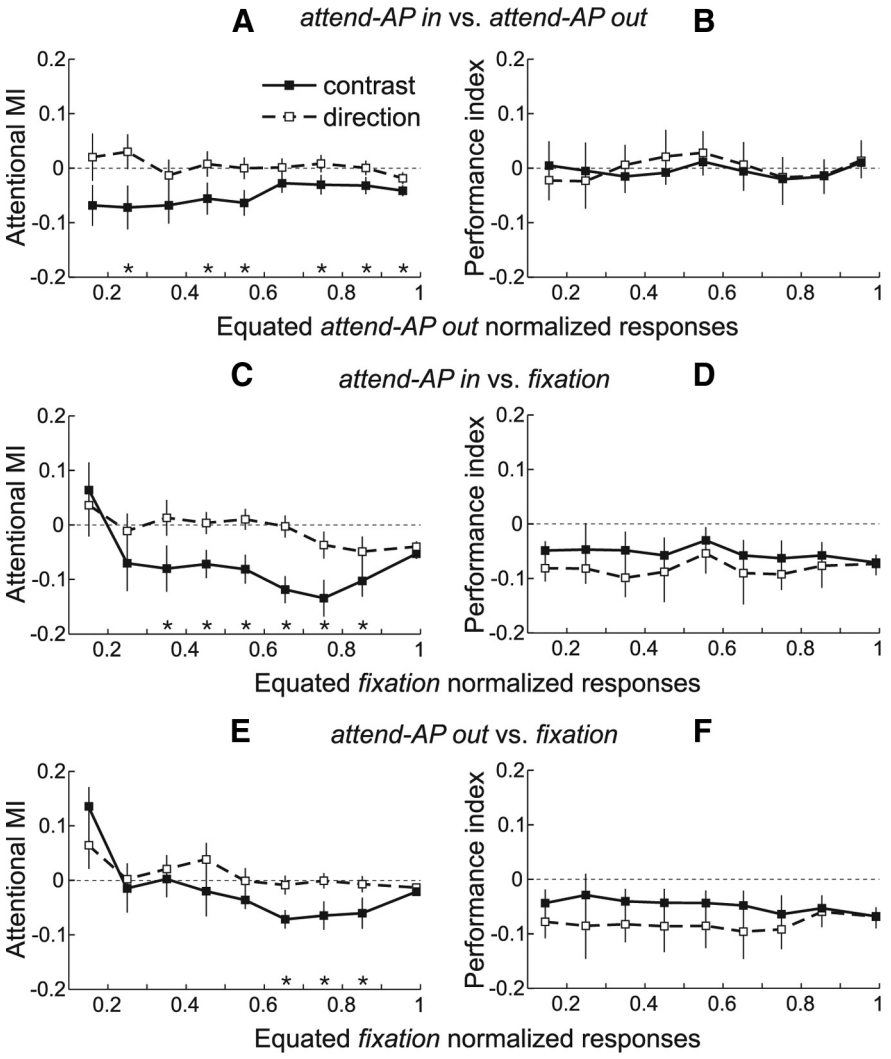


Figure 4. Attentional modulation of equated responses: population analysis. ***A***, Averaged population MIs between "attend-AP in" and "attend-AP out" trials across response levels equated during the "attend-AP out" condition. Error bars, SEM. Stars denote equated response levels with significant differences in MI between trials of the contrast (black squares) and direction (open squares) configuration. Note that in the population, the number of equated responses varied from cell to cell, and therefore the number of cells contributing to each response level also varied. ***B***, Averaged performance indices (perfMI) between "attend-AP in" and "attend-AP out" trials that contributed to each equated "attend-AP out" response level. Error bars, 95% confidence intervals (computed through a bootstrap procedure). ***C***, ***E***, Averaged population MIs between "attend-AP in" and "fixation" (***C***) and between "attend-AP out" and "fixation" trials (***E***) across responses equated during the "fixation" condition. ***D***, ***F***, Averaged performance indices between "fixation" trials and "attend-AP in" (***D***) or "attend-AP out" trials (***F***) that contributed to each equated "fixation" response level.

Figure 4*B* shows average perfMIs for trials corresponding to the response data shown in Figure 4*A*, and for the contrast and direction configuration. The perfMIs were not different from zero ($p > 0.2$, sign test for each stimulus pair), and were similar across all response levels and stimulus configurations ($p > 0.9$, Kruskal–Wallis ANOVA), indicating that the animals performed similarly regardless of the attended target location, and of the test-pattern contrast or direction. This suggests that the differences in response between the different conditions and stimulus configurations were not due to variations in the animals' attentional effort.

Figure 4*D* shows the perfMI corresponding to the response data in Figure 4*C*. The perfMIs were negative at each response level ($p < 0.05$, sign test), indicating that performance during "fixation" was higher than during "attend-AP in" trials. However, we did not find significant differences in the perfMI across response levels, and between trials of the contrast and direction configurations ($p > 0.2$, Kruskal–Wallis ANOVA). Additionally, perfMIs between trials of the "fixation" and "attend-AP out" conditions were very similar to the ones between trials of the "fixation" and "attend-AP in" conditions (Fig. 4, compare *D*, *F*) ($p > 0.3$, Kruskal–Wallis ANOVA). However, the response modulation between the conditions and the differences between contrast and direction pairs were larger in the latter case, reaching significance for almost all equated response levels but the two lowest and the highest (Fig. 4, compare *C*, *E*). We hypothesize that these larger effects are due to the combined effect of spatial and feature-based attention (attending to the antipreferred direction and inside the RF) in Figure 4*C* compared to the smaller effect of feature-

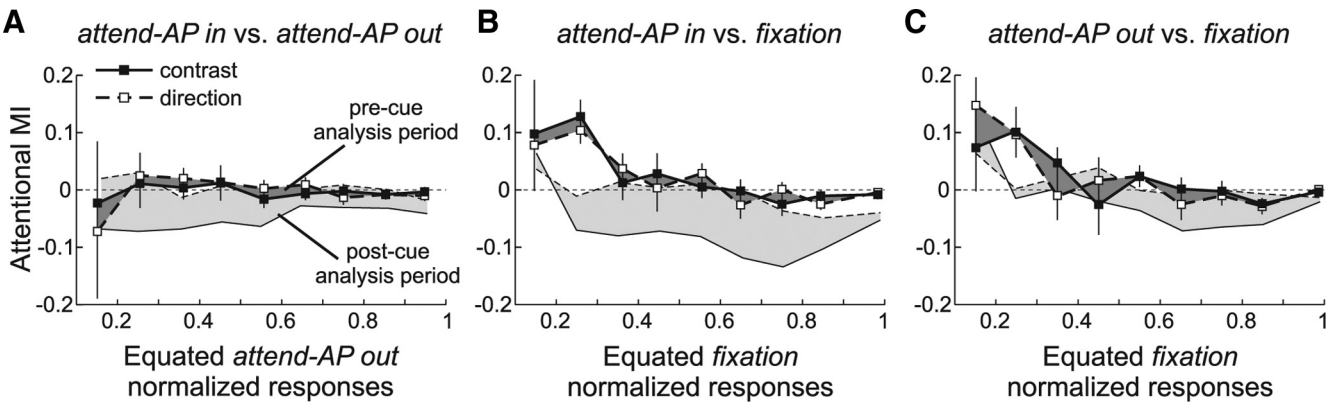


Figure 5. Control data. ***A–C***, MIs of the different comparisons for "attend-AP out" (***A***) and "fixation" (***B***, ***C***) responses equated during an analysis period before the attentional cue onset. The dark area represents the MI difference between the two stimulus configurations, computed during the precue analysis period. For comparison, the data from the postcue analysis period, as in Figure 4, are also shown (light area). The same conventions are used as in Figure 4.

based attention (attending to the antipreferred direction away from the RF) in Figure 4*E* (see Treue and Martínez Trujillo, 1999). These results strongly argue against the role of attentional effort in the differences in response modulation between the contrast and direction configurations.

We conducted an additional control to test whether our results were indeed caused by changes in the animals' allocation of attention following the instructions provided by the cue. We recomputed the MI across equated responses during a 400 ms time period previous to the cue onset. For all three comparisons, the differences in MIs between the contrast and direction data disappeared (Fig. 5, dark area) ($p > 0.05$, paired t test at each response level), indicating that effects isolated in the previous analysis were tightly coupled to the instructions provided by the cue.

All together, our findings argue against a stimulus-independent modulation of responses by attention. They rather show that contrast and direction pairs were differentially modulated by the three types of attention.

Attentional modulation as a function of the test-pattern contrast

Previous studies have reported that the magnitude of attentional modulation is stronger for neuronal responses evoked by stimuli with intermediate compared to low or high contrast (Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002). The previous analysis could not clearly replicate this result since it may have pooled, within a given response level, similar responses evoked by different stimulus pairs of the contrast configuration. To address this issue, we computed population responses ($n = 101$) in the different conditions for each contrast level of the test pattern.

We normalized each neuron's response in the contrast configuration to the response evoked by the maximal contrast stimulus pair (AP + preferred direction test pattern) during "fixation" trials, and then averaged responses to identical stimulus pairs in each condition across units (Fig. 6*A*). Responses in the "fixation" condition (gray) decreased as the contrast of the test pattern decreased. Attending to the AP pattern yielded an overall response suppression that varied as a function of contrast (larger along the slope of the contrast response function and smaller on the extremes). This suppression was stronger in the "attend-AP in" (black) relative to the "attend-AP out" condition (dashed line). We quantified these observations by computing MIs for the different comparisons (Fig. 1*A*) for each individual neuron and stimulus combination, and then averaging across indices.

Figure 6*B* shows the average MIs for the comparison "attend-AP in" versus "attend-AP out" conditions. For all stimulus combinations the MIs were significantly lower than zero (open squares, $p < 0.05$, t test), indicating that responses were suppressed in the former condition. Responses during trials with the highest-contrast test pattern were suppressed by 5.5% (rightmost data point, MI $= -0.027$, $p < 0.02$, t test), and reached 15% suppression as the contrast of the test pattern decreased (MI $= -0.083$, $p < 0.0001$, t test; second data point from left to right). A paired comparison between MIs for high- and intermediate-contrast stimuli yielded significance ($p < 0.005$, paired t test). These results indicate that the response modulation produced by shifting spatial attention into the RF was stronger for intermediate-contrast stimuli, and decreased for the lowest- and highest-contrast stimuli. These results are similar to the ones reported in previous studies of attention (Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002).

To examine whether the changes in the strength of attentional modulation across the various contrasts was due to changes in task difficulty, we computed the perfMI between trials of the "attend-AP in" and "attend-AP out" conditions (Fig. 6*B*, dashed line). Performance in both conditions was similar (perfMI not different from zero, $p > 0.3$, sign test for each contrast pair) and did not change across stimuli ($p > 0.8$, Kruskal–Wallis ANOVA). This result demonstrates that all trial types in both tasks were correspondingly difficult, suggesting that this variable could not have caused the observed pattern of attentional modulation. This conclusion also holds against the improbable argument that our performance measurements may not reflect attentional effort since the trials that theoretically would lead to the highest and lowest efforts (100% vs ~0% contrast, rightmost and leftmost data points, respectively) led to a similar attentional modulation, i.e., the magnitude of the attentional modulation corresponding to these trials was weaker and more similar than the one corresponding to trials with intermediate contrast ($p < 0.005$, paired t test).

We further compared the responses between the "attend-AP in" and "fixation" conditions (Fig. 6*C*). Except for the lowest-contrast combination (black square), the MI was significantly below zero (open squares, $p < 10^{-6}$, t test), indicating that responses were suppressed in the "attend-AP in" condition relative to "fixation." Responses during trials with the highest-contrast stimulus combination were suppressed by 10% (rightmost data point, MI $= -0.051$, $p < 10^{-6}$, t test), and reached 23% suppres-

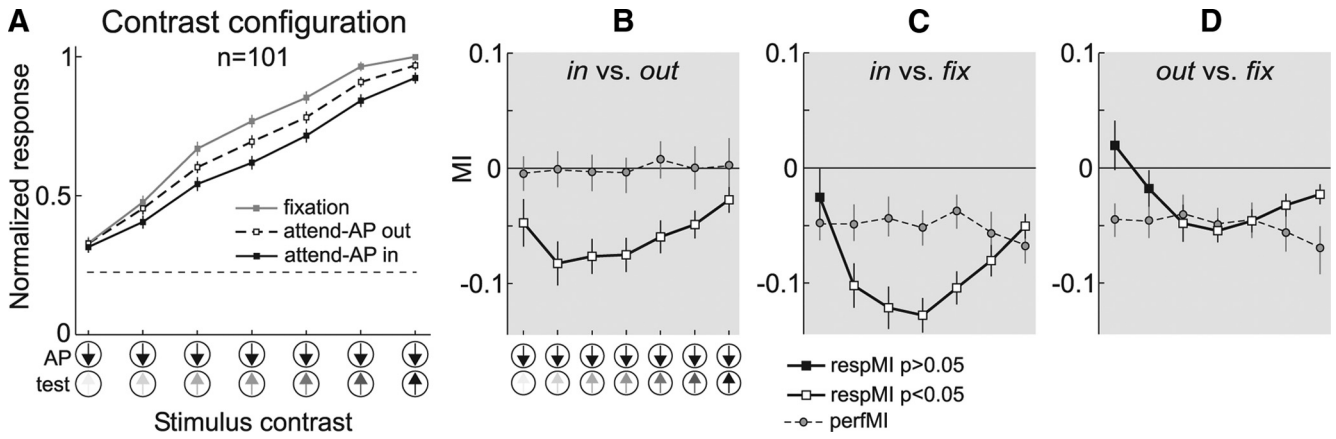


Figure 6. Attentional modulation in the contrast configuration. ***A***, Averaged normalized population responses ($n = 101$) for each stimulus pair of the contrast configuration, in the "fixation" (gray), "attend-AP in" (black), and "attend-AP out" (dashed) conditions. The horizontal line indicates spontaneous response without stimulus. Error bars, SEM. ***B–D***, Averaged response MI (respMI, solid line and squares) and performance MI (perfMI, dashed line and circles) in the contrast configuration between the "attend-AP in" and "attend-AP out" (***B***), "attend-AP in" and "fixation" (***C***), and "attend-AP out" and "fixation" conditions (***D***). The respMIs were computed for each neuron and then pooled across units ($n = 101$). Similarly, the perfMIs were computed for each recording session and then pooled across sessions ($n = 96$). Open squares denote respMIs significantly different from zero and black squares denote nonsignificant ones. Error bars, SEM for the respMI and 95% confidence intervals for the perfMI.

sion during trials with intermediate contrast (MI $= -0.128$, $p < 10^{-9}$, t test). The difference between these two points (13%) was highly significant ($p < 10^{-6}$, paired t test). Relative to the previous comparison, these attentional effects are stronger, likely because they combine the effects of spatial and feature-based attention (i.e., directing attention from fixation to the antipreferred direction inside the RF). Although the animals' performance was higher during "fixation" compared to "attend-AP in" trials across all stimulus combinations (Fig. 6*C*, dashed line) (perfMI lower than zero, $p < 10^{-6}$, sign test), these differences do not follow the shape of the differences in response quantified by the attentional MIs. The perfMI was similar across all contrast combinations ($p > 0.1$, Kruskal–Wallis ANOVA). Again, in this dataset, the differences in performance cannot account for the shape of the differences in attentional modulation (dashed line vs solid line).

We finally compared responses between the "fixation" and "attend-AP out" conditions (Fig. 6*D*). Here, the magnitude of the attentional modulation was smaller than in the previous comparison, but it followed a similar shape. The MIs were significantly different from zero across most contrast levels (white data points, $p < 0.02$, t test), and the pattern of larger differences for intermediate-contrast stimuli was also present. Responses to intermediate contrast were significantly suppressed by 10.5% (MI $= -0.055$) compared to 4.5% (MI $= -0.023$) to the highest-contrast stimulus ($p < 0.01$, paired t test). The perfMI did not follow the trend observed in the MIs, confirming that the pattern of attentional modulation was not due to differences in performance between conditions. Noticeably, attentional modulation in Figure 6*C* was clearly larger than in Figure 6*D*, but the perfMIs were almost identical. This may reflect in the former case the combined effects of feature- and space-based attention, while in the latter case the isolated effect of feature-based attention.

In summary, these data show that spatial and feature-based attention and their combination modulate responses to intermediate-contrast stimuli more strongly than responses to low- and high-contrast stimuli. This demonstrates the contrast dependency of the modulation.

Attentional modulation as a function of the test-pattern direction

Our data in the direction configuration allowed us to further examine whether the magnitude of the attentional modulation changed as a function of the test-pattern motion direction in the absence of changes in stimulus contrast. We computed averaged population responses during trials of the direction configuration in the three behavioral conditions (Fig. 7*A*). Surprisingly, we found that relative to "fixation" (gray), responses in the "attend-AP in" (black) and "attend-AP out" (dashed line) conditions were suppressed for directions of the test pattern similar to the preferred (rightmost data points) but enhanced for directions progressively closer to the attended antipreferred direction (leftmost data points). This suppression/enhancement pattern was also consistent across the majority of the individual neurons (see cell example in Fig. 3*A*, right).

We quantitatively assessed the magnitude of these effects by determining for each neuron and direction combination the MIs between the "fixation" and the "attend-AP in" (Fig. 7*C*) or "attend-AP out" (Fig. 7*D*) conditions. In the first comparison the average MIs pooled across neurons became more positive as the direction of the test pattern approached the attended direction. "Attend-AP in" responses were modulated within a 23.5% range, shifting from 10% response suppression (rightmost data point, MI $= -0.051$, $p < 10^{-6}$, t test) to 13.5% enhancement (leftmost data point, MI $= 0.063$, $p < 0.01$, t test) relative to the "fixation" responses. The results for the comparison "fixation" versus "attend-AP out" (Fig. 7*D*) (which mainly isolates the effect of feature-based attention) were similar. Here, attentional effects shifted from 4.5% response suppression (MI $= -0.023$, $p < 0.02$, t test) to 16% enhancement (MI $= 0.074$, $p < 0.005$, t test). This shift in MIs from negative to positive in both comparisons, cannot be explained by variations in performance, which was higher in the "fixation" compared to the other two conditions for each stimulus combination (negative perfMI, dashed line, $p < 10^{-9}$, sign test), but constant across all stimuli ($p > 0.1$, Kruskal–Wallis ANOVA).

One possible explanation for this effect is that in our task, feature-based attention (Martinez-Trujillo and Treue, 2004) dif-

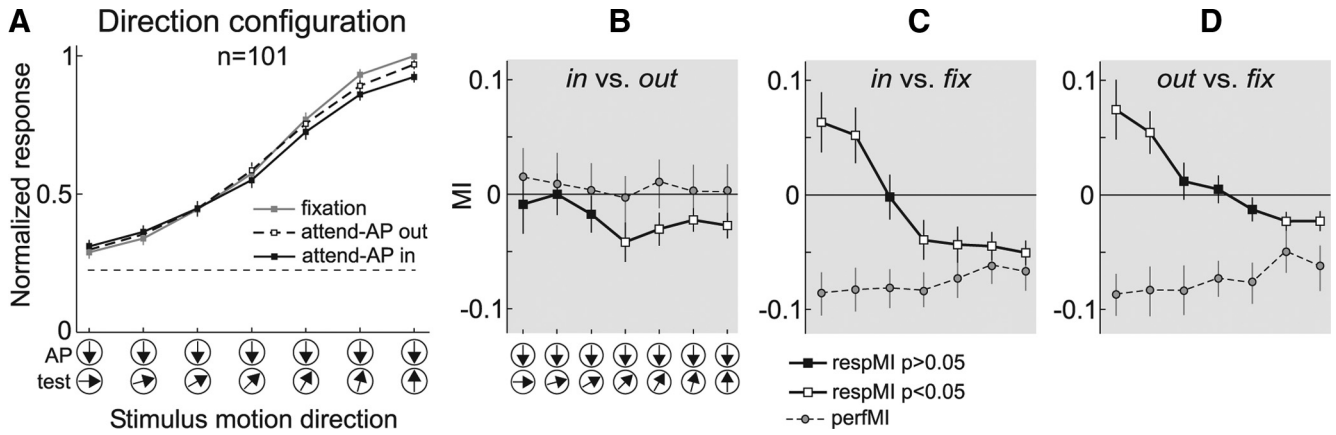


Figure 7. Attentional modulation in the direction configuration. ***A***, Averaged normalized population responses ($n = 101$) in the direction configuration. ***B–D***, Averaged respMI and perfMI for the three comparisons. The same conventions are used as in Figure 6.

ferentially modulated the activation strength of two populations of direction-selective inputs carrying signals from each pattern into the recorded MT neuron. Several models have proposed this input modulation as a possible mechanism of attention (Reynolds et al., 1999; Ghose and Maunsell, 2008; Ghose, 2009; Reynolds and Heeger, 2009). When the animal allocates attention to the MT neuron's antipreferred direction (AP pattern) the strength of the inputs activated by that pattern (and therefore selective for that direction) would be enhanced. On the other hand, the strength of the inputs activated by the test pattern could change from a relatively strong suppression, when the activated inputs are selective for a direction 180° away from the attended direction, to a progressively weaker suppression when the activated inputs are selective for directions progressively away from the preferred and closer to the attended antipreferred direction. When the total amount of enhancement of inputs activated by the AP pattern overcomes the progressively weaker suppression of inputs activated by the test pattern, the modulation switches sign, becoming predominantly a response enhancement relative to fixation.

Note that within this interpretation what defines the enhancement/suppression pattern due to feature-based attention is the relationship between the selectivity of the inputs and the attended direction. The same feature-based effects defined with respect to the MT neuron selectivity will show no changes in response since in our experiments the animals attended to the same motion direction across the different direction combinations (see Martinez-Trujillo and Treue, 2004). Also note that this effect of feature-based attention is defined relative to the "fixation" (neutral with respect to the attended feature) condition.

Although we cannot directly test this hypothesis, we could test at least one of its predictions. Removing the contribution of feature-based attention should abolish the shift in response modulation. We therefore compared responses between "attend-AP in" and "attend-AP out" trials, in which the modulation should mainly reflect the effect of spatial attention. We found that the shift in modulation disappeared (Fig. 7*B*). The response increase for directions of the test pattern more similar to the attended direction vanished (black data points), while a small response suppression remained for directions more dissimilar to the attended direction (white data points). However, across different direction pairs, there were no significant differences ($p > 0.05$, one-way ANOVA). This effect was not due to changes in the animal performance, since performance was comparable in all trial types of both conditions (perfMI, dashed line, $p > 0.8$, Kruskal–Wallis ANOVA). Overall, these data are compatible with the effect of feature-based attention on the strength of direction-selective inputs carrying signals from the AP and test pattern into the recorded neurons.

A question that arises when comparing the results of Figure 7 and Figure 4, *A*, *C*, and *E*, is why the attentional modulation as a function of motion direction was, in the latter case, almost negligible. The explanation is that in this analysis we pooled, within a given response level, responses to stimuli with different directions. As shown in Figure 7, *C* and *D*, the modulation changed as a function of the test-pattern direction, thus, pooling responses corresponding to different directions would "cancel out" or "attenuate" attentional effects.

Discussion

We showed that, first, attentional modulation of similar response levels was systematically dependent on the visual stimulus. Second, attentional effects were stronger for intermediate-contrast stimuli. Third, the sign and magnitude of attentional modulation changed with the test-pattern motion direction, an effect attributable to feature-based attention.

Attentional modulation as a function of response level

Previous studies have reported attentional modulation of responses that are independent of stimulus orientation, direction, and contrast in MT and V4 neurons (McAdams and Maunsell, 1999; Treue and Martínez Trujillo, 1999; Williford and Maunsell, 2006), and a recent study has reported a contrast-independent modulation of V1 responses (Thiele et al., 2009). Others, however, have reported contrast-dependent modulation in areas MT and V4 (Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002). Here, we found that attentional modulation in MT neurons was stronger for the contrast relative to the direction stimuli. Unlike previous studies, our approach produced a situation where distinct stimuli evoked the same response level. When controlling for the allocation of attention, this situation yields a clear prediction for the stimulus-independent modulation, i.e., a similar amount of response change for both stimulus configurations. Our data clearly argue against this prediction.

Furthermore, spatial, feature-based attention, and their combination produced similar effects, suggesting that the different types of attention share similar constraints. Differences between the contrast and direction data were also larger when combining

spatial and feature-based attention than when each one acted alone, supporting previous reports of additive interactions between these types of attention (Treue and Martínez Trujillo, 1999; McAdams and Maunsell, 2000; Hayden and Gallant, 2009).

Modulation as a function of stimulus contrast

Our results in the contrast configuration agree with previous findings of a stronger attentional modulation for intermediate-contrast stimuli (Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002); however, they apparently disagree with other reports of contrast-independent modulations (Williford and Maunsell, 2006; Thiele et al., 2009).

One may argue that differences in attentional effort associated with contrast changes may have confounded our results (Spitzer et al., 1988; Ghose and Maunsell, 2002; Boudreau et al., 2006), therefore explaining differences between studies. In our design, attentional effort was kept similar across both configurations by instructing the animals to always direct attention to a high-contrast stimulus (see performance data). It may be possible that when the test pattern had lower contrast the task was somewhat easier, and that although that was not reflected in the animals' performance, it may have influenced our results. At least three other findings discard this argument. First, for the lowest- and highest-contrast stimuli, where this hypothesis predicts the largest differences, the response modulation was similar, but it was in both cases different from the one for intermediate-contrast stimuli. Second, the comparisons "fixation" versus either "attend-AP in," or "attend-AP out," yielded considerable differences in modulation strength, however, their corresponding perfMIs were identical. Third, attentional effects were larger in the contrast compared to the direction configuration. The effect of task difficulty would have predicted the opposite result (i.e., larger modulation in the direction data where contrast was always high).

A more likely source of variability in results among studies is differences in task timing. The study reporting the largest variability in results across neurons has used the shortest stimulus presentation (94 ms) and analysis period (150 ms) (Williford and Maunsell, 2006). Another V4 study used a longer stimulus presentation (250 ms) and analysis period (400 ms) (Reynolds et al., 2000), obtaining a contrast-dependent modulation similarly as in our study (see also Martínez-Trujillo and Treue, 2002). This suggest that the dependency of the modulation on contrast becomes more evident for longer stimulus presentations and analysis periods. At least two different factors may cause this effect. First, because response latency in visual neurons increases when decreasing contrast (Reynolds and Desimone, 2003; Lee et al., 2007), integrating responses during early and short periods may underestimate firing rates and attentional effects associated to low-contrast stimuli. Second, transient stimuli may initially trigger exogenous attention, which if interacting with endogenous attention may produce different effects on responses during early periods. Interestingly, psychophysical experiments in humans have reported contrast-independent modulation of psychometric functions by exogenous attention, and contrast-dependent modulation by endogenous attention (Carrasco, 2006).

A number of studies have proposed quantitative models in which attention changes inputs' strength to an area normalization circuit (Reynolds et al., 1999; Raizada and Grossberg, 2003; Reynolds and Chelazzi, 2004; Ghose, 2009; Lee and Maunsell, 2009; Reynolds and Heeger, 2009). Our results are compatible with these models. On the other hand, they are not compatible with models in which a neuron's response function is scaled or shifted along the response axis by attention (Treue and Martínez Trujillo, 1999; Williford and Maunsell, 2006; Boynton, 2009; Thiele et al., 2009).

An intriguing question is why the only existing study in V1 neurons reported a contrast-independent attentional modulation (Thiele et al., 2009). One possible explanation is that at the level of V1 neurons, attention uses a different mechanism than input strength modulation, wherein contrast increases do not constraint the modulation's magnitude. At least for feature-based attention this may make sense, since selectivity for some stimulus features such as motion direction arises in V1 (Peterson et al., 2004). Thus, an input modulation based on the attended feature (see next section) may not be possible. A target question for future studies is whether attentional modulation in thalamic lateral geniculate nucleus (McAlonan et al., 2008), and superior colliculus (Ignashchenkova et al., 2004) neurons, which are not direction selective, is contrast independent.

One may ask whether our results only apply to cases in which the effect of attention is associated with a decrease in firing rate (Martínez-Trujillo and Treue, 2002) rather than with an increase (Reynolds et al., 2000; Williford and Maunsell, 2006; Thiele et al., 2009). In our experimental design, these two processes likely occur in two different neuronal populations selective to opposite features (i.e., motion direction in MT, orientation in V4). We are inclined to believe that both process are constrained in contrast, because in the study of Reynolds et al. (1999), and in some of the neurons recorded by Williford and Maunsell (2006), increases in response by attention showed a similar contrast dependency as the response decreases shown in our study. However, whether the strength of the contrast-dependent attentional modulation between these two processes varies remains to be clarified.

Modulation as a function of the test-pattern direction

A surprising finding in our study was that in the direction configuration when animals switched attention between the fixation spot and the AP pattern inside or outside the RF, responses to the stimulus pair changed as a function of the test-pattern direction (Fig. 7). This result is different from the quasi-multiplicative modulation previously reported using a similar task (Treue and Martínez Trujillo, 1999). We believe there are two main reasons for these differences. First, that study sampled a limited number of test-pattern directions (every 30°), while here we sampled the direction space every 15°, providing a better resolution to detect differences in attentional modulation along the response function. Second, in the current study we may have included a sample of MT neurons more narrowly tuned for motion direction than in the previous study, since we actively searched for neurons that showed a similar decay in response within a 90° range as with decreases in contrast. In the previous study, we recorded from any neuron that showed direction selectivity within a 360° range, which may have resulted in a sample of more broadly tuned neurons where the direction-dependent modulation may become less evident.

We hypothesize that the test-pattern direction-dependent modulation may be caused by feature-based attention differentially modulating the strength of direction-selective inputs carrying signals from the AP and test pattern into the recorded neurons (Reynolds et al., 1999). These inputs may originate from area V1 direction-selective cells projecting toward the recorded MT neuron (Born and Bradley, 2005; Rust et al., 2006), whose responses could be modulated by attention (Motter, 1993; Roelfsema et al., 1998; Khayat et al., 2006). We did not directly test this hypothesis; however, two findings support it. First, the shift from suppression to enhancement was present when the animals

switched attention between the fixation spot and the AP pattern outside the RF (feature-based attention). In the contrast configuration the shift was not present, indicating that it was not due to a bias introduced by low firing rates when the test-pattern direction deviated from the preferred direction; for low test-pattern contrasts, responses were also low. Second, the response shift disappeared when the animal switched attention from the AP pattern outside the RF to the one inside (spatial attention).

In general, our results support models in which spatial and feature-based attention modulate the strength of direction-selective inputs from attended and unattended stimuli into the MT normalization circuit. Clarifying whether this modulation occurs at the level of population of neurons in areas projecting toward MT, or at the level of synaptic inputs into MT neurons, or both, remains as a challenge for future studies.

References

Born RT, Bradley DC (2005) Structure and function of visual area MT. Annu Rev Neurosci 28:157–189.

Boudreau CE, Williford TH, Maunsell JHR (2006) Effects of task difficulty and target likelihood in area V4 of macaque monkeys. J Neurophysiol 96:2377–2387.

Boynton GM (2009) A framework for describing the effects of attention on visual responses. Vision Res 49:1129–1143.

Buracas GT, Boynton GM (2007) The effect of spatial attention on contrast response functions in human visual cortex. J Neurosci 27:93–97.

Busse L, Katzner S, Treue S (2008) Temporal dynamics of neuronal modulation during exogenous and endogenous shifts of visual attention in macaque area MT. Proc Natl Acad Sci U S A 105:16380–16385.

Carrasco M (2006) Covert attention increases contrast sensitivity: psychophysical, neurophysiological and neuroimaging studies. Prog Brain Res 154:33–70.

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

Ekstrom LB, Roelfsema PR, Arsenault JT, Kolster H, Vanduffel W (2009) Modulation of the contrast response function by electrical microstimulation of the macaque frontal eye field. J Neurosci 29:10683–10694.

Ghose GM (2009) Attentional modulation of visual responses by flexible input gain. J Neurophysiol 101:2089–2106.

Ghose GM, Maunsell JHR (2002) Attentional modulation in visual cortex depends on task timing. Nature 419:616–620.

Ghose GM, Maunsell JHR (2008) Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4. J Neurosci 28:5115–5126.

Hayden BY, Gallant JL (2009) Combined effects of spatial and feature-based attention on responses of V4 neurons. Vision Res 49:1182–1187.

Ignashchenkova A, Dicke PW, Haarmeier T, Thier P (2004) Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. Nat Neurosci 7:56–64.

Khayat PS, Spekreijse H, Roelfsema PR (2006) Attention lights up new object representations before the old ones fade away. J Neurosci 26:138–142.

Lee J, Maunsell JH (2009) A normalization model of attentional modulation of single unit responses. PLoS One 4:e4651.

Lee J, Williford T, Maunsell JH (2007) Spatial attention and the latency of neuronal responses in macaque area V4. J Neurosci 27:9632–9637.

Li X, Lu ZL, Tjan BS, Dosher BA, Chu W (2008) Blood oxygenation level-dependent contrast response functions identify mechanisms of covert attention in early visual areas. Proc Natl Acad Sci U S A 105:6202–6207.

Martínez-Trujillo J, Treue S (2002) Attentional modulation strength in cortical area MT depends on stimulus contrast. Neuron 35:365–370.

Martinez-Trujillo JC, Treue S (2004) Feature-based attention increases the selectivity of population responses in primate visual cortex. Curr Biol 14:744–751.

Maunsell JH, Treue S (2006) Feature-based attention in visual cortex. Trends Neurosci 29:317–322.

McAdams CJ, Maunsell JH (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. J Neurosci 19:431–441.

McAdams CJ, Maunsell JH (2000) Attention to both space and feature modulates neuronal responses in macaque area V4. J Neurophysiol 83:1751–1755.

McAlonan K, Cavanaugh J, Wurtz RH (2008) Guarding the gateway to cortex with attention in visual thalamus. Nature 456:391–394.

Motter BC (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. J Neurophysiol 70:909–919.

Moulden B, Kingdom F, Gatley LF (1990) The standard deviation of luminance as a metric for contrast in random-dot images. Perception 19:79–101.

Murray SO (2008) The effects of spatial attention in early human visual cortex are stimulus independent. J Vis 8:2.1–2.11.

Peterson MR, Li B, Freeman RD (2004) The derivation of direction selectivity in striate cortex. J Neurosci 24:3583–3591.

Raizada RD, Grossberg S (2003) Towards a theory of the laminar architecture of cerebral cortex: computational clues from the visual system. Cereb Cortex 13:100–113.

Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. Annu Rev Neurosci 27:611–647.

Reynolds JH, Desimone R (2003) Interacting roles of attention and visual salience in V4. Neuron 37:853–863.

Reynolds JH, Heeger DJ (2009) The normalization model of attention. Neuron 61:168–185.

Reynolds JH, Chelazzi L, Desimone R (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci 19:1736–1753.

Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. Neuron 26:703–714.

Roelfsema PR, Lamme VAF, Spekreijse H (1998) Object-based attention in the primary visual cortex of the macaque monkey. Nature 395:376–381.

Rust NC, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. Nat Neurosci 9:1421–1431.

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

Thiele A, Pooresmaeili A, Delicato LS, Herrero JL, Roelfsema PR (2009) Additive effects of attention and stimulus contrast in primary visual cortex. Cereb Cortex 19:2970–2981.

Treue S (2001) Neural correlates of attention in primate visual cortex. Trends Neurosci 24:295–300.

Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399:575–579.

Williford T, Maunsell JHR (2006) Effects of spatial attention on contrast response functions in macaque area V4. J Neurophysiol 96:40–54.

2.4 Frequency-Dependent Attentional Modulation of Local Field Potential Signals in Macaque Area MT

During extracellular recordings, an electrode positioned close to a neuron registers the train of action potentials produced by the cell. This spiking activity is easily distinguishable from other signals in the extracellular environment due to its high frequency (usually above 300Hz). In the lower frequencies (below 120Hz), another physiological signal, the local field potential (LFP), can be recorded from the same electrode. The LFP is thought to represent the sum of all local currents within the recording area covered by the electrode tip. Such currents are likely the result of the contribution of inputs into the area as well as its local intracortical processing. Unlike spikes, LFPs are characterized by slow waveforms, which are traditionally classified in different frequency bands.

Previous studies have demonstrated that attention modulates the spiking activity of visual neurons. However, little is known about the effects of attention on the strength of LFPs oscillations (power). To address this issue, we recorded LFPs together with neuronal spiking activity in area MT in two monkeys while they performed an attentional task.

We found that LFPs in higher frequency bands (24 to 164 Hz) were similarly tuned as the spiking activity for changes in the motion direction and contrast of moving random dot patterns (RDPs). However, the tuning was absent in low frequency bands (< 24 Hz). Attending to one of two moving RDPs inside a neurons receptive field either enhanced or suppressed the LFP power depending on the features of the second, unattended stimulus. Again, this modulation was similar to that of the neurons spiking activity. In the low frequency band, however, attention always suppressed the LFP power, independently of the unattended stimulus' features.

These results demonstrate that high- and low-frequency LFP oscillations carry different information about a visual stimulus. Furthermore, they show that although both are modulated by attention, the strength of the modulation is different in the different frequency bands.

Behavioral/Systems/Cognitive

Frequency-Dependent Attentional Modulation of Local Field Potential Signals in Macaque Area MT

Paul S. Khayat,[1,2] Robert Niebergall,[1,3] and Julio C. Martinez-Trujillo[1,2]

[1]Cognitive Neurophysiology Laboratory, Department of Physiology, McGill University, Montréal, Québec H3G 1Y6, Canada, [2]Research Center in Neuropsychology and Cognition, University of Montreal, Montréal, Québec H3C 3J7, Canada, and [3]Cognitive Neuroscience Laboratory, German Primate Center, 37077 Goettingen, Germany

Visual attention modulates neuronal responses in primate motion processing area MT. However, whether it modulates the strength local field potentials (LFP-power) within this area remains unexplored, as well as how this modulation relates to the one of the neurons' response. We investigated these issues by simultaneously recording LFPs and neuronal responses evoked by moving random dot patterns of varying direction and contrast in area MT of two male monkeys (*Macaca mulatta*) during different behavioral conditions. We found that: (1) LFP-power in the γ (30–120 Hz), but not in the δ (2–4 Hz), θ (4–8 Hz), α (8–12 Hz), β_1 (12–20 Hz), and β_2 (20–30 Hz) frequency bands, was tuned for motion direction and contrast, similarly to the neurons' response, (2) shifting attention into a neuron's receptive field (RF) decreased LFP-power in the bands below 30 Hz (except the θ band), whereas shifting attention to a stimulus motion direction outside the RF had no effect in these bands, (3) LFP-power in the γ band, however, exhibited both spatial- and motion direction-dependent attentional modulation (increase or decrease), which was highly correlated with the modulation of the neurons' response. These results demonstrate that in area MT, shifting attention into the RFs of neurons in the vicinity of the recording electrode, or to the direction of a moving stimulus located far away from these RFs, distinctively modulates LFP-power in the various frequency bands. They further suggest differences in the neural mechanisms underlying these types of attentional modulation of visual processing.

Introduction

The frequency of action potentials fired by neurons in area middle temporal (MT) of monkeys encodes the contrast, direction, and speed of moving stimuli (Zeki, 1980; Felleman and Kaas, 1984; Sclar et al., 1990). The firing rate of these neurons is also modulated when attention is directed into their receptive fields (RFs), or to the motion direction of a stimulus outside the RF (Treue and Maunsell, 1996; Seidemann and Newsome, 1999; Treue and Martínez Trujillo, 1999). A previous study has demonstrated that the amplitude of local field potentials oscillations (LFP-power) in this area encodes, within certain frequencies, motion direction and speed (Liu and Newsome, 2006). However, it remains uninvestigated whether directing attention to an object's spatial position, or feature encoded by MT neurons, modulates LFP-power in this area.

LFPs contain oscillations with frequencies below 200 Hz that can be recorded from the same electrode as higher frequency spikes fired by single units. LFPs are thought to represent synaptic activity within a local network (in the recorded cell's vicinity), as well as voltage-dependent membrane oscillations, spike components and afterpotentials, and inputs from other brain regions (Mitzdorf, 1985, 1987; Kruse and Eckhorn, 1996; Buzsáki, 2002; Logothetis, 2003; Logothetis and Wandell, 2004; Katzner et al., 2009; Khawaja et al., 2009). Understanding how LFP-power in area MT relates to changes in stimulus attributes such as contrast and direction, and how it is influenced by visual attention may reveal important aspects of the computations underlying sensory and cognitive processing in the primate brain.

In other visual areas than MT, previous studies have reported that LFP-power in the γ frequencies (>25 Hz) is selective to stimulus attributes (Frien et al., 2000; Kayser and König, 2004; Henrie and Shapley, 2005; Berens et al., 2008; Katzner et al., 2009). In area V4, spatial attention enhances γ-band power, while it generally decreases power in lower frequencies (Fries et al., 2001, 2008; Taylor et al., 2005). In area V1, however, Lakatos et al. (2008) reported that during certain tasks attention could increase LFP-power in the δ band (<4 Hz) (see also Schroeder and Lakatos, 2009). These attentional effects in different bands may be a generalized finding across visual cortical areas that reflects the local and/or global neural computations underlying attentional filtering of behaviorally relevant signals (Fries, 2009). If this is the case, we should find frequency-dependent modulations of LFPs recorded from area MT during attentional tasks.

Here, we explored these issues by measuring LFP-power and spiking activity evoked by moving random dot patterns with various directions and contrasts in area MT of macaques during different experimental conditions. First, we determined the direction and contrast selectivity of the LFPs in different frequency bands, as well as the correlation of the LFP-power in each band with the neurons spiking activity. Second, we isolated in each

Received Jan. 24, 2010; revised March 30, 2010; accepted April 9, 2010.

This study was supported by a Natural Sciences and Engineering Research Council of Canada Fellowship awarded to P.S.K., and Canada Foundation for Innovation and Canadian Institutes of Health Research grants awarded to J.C.M.-T. We thank W. Kucharski and S. Nuara for technical assistance.

Correspondence should be addressed to Paul S. Khayat, Department of Physiology, McGill University, Room 1220, 3655 Prom. Sir W. Osler, Montréal, QC H3G 1Y6, Canada. E-mail: paul.khayat@mcgill.ca.

DOI:10.1523/JNEUROSCI.0404-10.2010

 0270-6474/10/307037-12$15.00/0

band the effects of attending to an object's spatial position (inside vs outside a neuron's RF), or to a feature encoded by MT neurons (motion direction), by contrasting measurements corresponding to the different conditions.

Materials and Methods

Two male macaque monkeys (*Macaca mulatta*) participated in the experiments. All procedures complied with the Canadian Council of Animal Care guidelines and were approved by the McGill University animal care committee. Standard surgical and electrophysiological techniques were used to record neuronal activity in area MT (Khayat et al., 2010).

Behavioral task. On each trial, the animal had to press a button and fixate within a circular window of 1.5° diameter centered on a small fixation spot (0.06° square). After 470 ms, two pairs of moving RDPs appeared, one located inside the RF of the recorded MT neuron, and the other located outside, in the opposite hemifield (Fig. 1). Each pair consisted of a high contrast RDP moving in the neuron's antipreferred direction (AP-pattern) and a test RDP (test-pattern) that: (1) had the same contrast as the AP-pattern but moved, from trial to trial, in different directions (direction configuration), or (2) moved in the neuron's preferred direction but had, from trial to trial, different contrasts (contrast configuration).

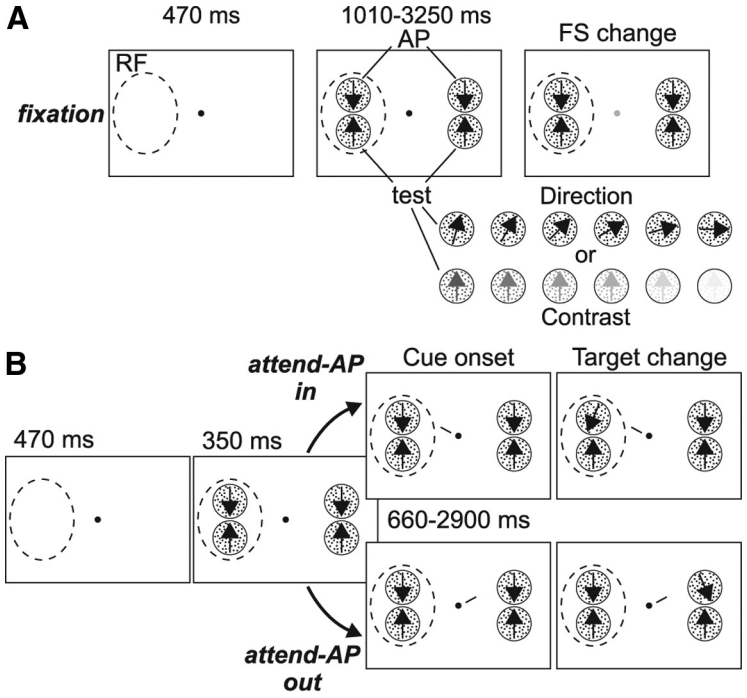


Figure 1. Behavioral task. ***A***, Sequence of events during fixation trials: The monkey had to fixate a central spot and respond to a brief change in its luminance, while ignoring the RDPs. The dashed circle represents the neuron's RF. From trial to trial, the AP pattern was paired with a test-pattern that had the same contrast but moved in different directions, or moved in the neuron's preferred direction (upward arrow) but had different contrast levels (see sketches at the bottom). ***B***, Sequence of events during attend-AP trials. The monkey was cued to covertly attend to the AP-pattern located inside (Attend-AP in condition, top) or outside (Attend-AP out condition, bottom) the RF, and respond when the cued pattern briefly changed direction. From trial to trial, the irrelevant test-patterns could have different motion directions or contrast levels (as shown in ***A***) (see also Materials and Methods).

Three different task conditions were used. In the fixation condition (Fig. 1*A*), the animal detected a subtle luminance change in the fixation spot, which occurred at a random interval between 1010 and 3250 ms after stimulus onset. Here, the animal was required to ignore changes in the direction of any of the RDPs. In the other conditions (Fig. 1*B*), a small line (1° length) appeared next to the fixation spot 350 ms after stimulus onset. This cue-line pointed toward one of the AP-patterns, thereby instructing the monkey to direct attention to this pattern located either inside (attend-AP in), or outside (attend-AP out) the RF. After a variable delay of 660–2900 ms from cue onset, the target (i.e., cued AP-pattern) underwent a brief direction change (23° during 100 ms). The animal had to release the button within a response time interval of 150–500 ms after the change to receive a juice reward.

In the two attentional conditions of this task (attend-AP in/out), attention is always on a high contrast pattern moving in the antipreferred direction (AP-pattern). This avoided the possibility that the animal's performance would substantially change while manipulating the direction or contrast of the test-pattern across trials. Also, by using a combination of antipreferred + test-pattern, we could obtain a modulation of the firing rate that ranges from very low (i.e., close to spontaneous response evoked by the combination of the AP-pattern with the 90° away direction, or lowest contrast test-pattern) to ~70–80% of the response to the preferred direction (elicited by the combination of AP-pattern and preferred test-pattern direction) (Martínez-Trujillo and Treue, 2002; Khayat et al., 2010). This allows us to explore direction and contrast tuning as well as the effect of attention within this 70–80% response range.

To ensure that during the attend-AP in and out conditions the monkey was focusing attention on the target, on half of the corresponding trials the uncued AP-pattern (the distracter), located in the opposite visual hemifield, briefly changed direction. The monkey had to ignore this distracter change and wait until the target changed. Trials in which the monkey responded to the distracter change, or broke fixation before the target's change occurred were terminated without reward and considered errors. The different trial types were presented in random sequence, and both animals performed between 6 and 15 trials (median = 12 trials) per stimulus type in each behavioral condition. Only correctly performed trials, with no change events within the analyzed period, were included in the analysis.

By comparing responses between the different conditions, we could isolate the effects of: (1) spatial attention (attend-AP out vs attend-AP in: directing attention from the AP-pattern outside the RF to the AP-pattern inside). Here, the potential target stimuli (AP-patterns inside and outside the RF) differ in their spatial positions, but they share the same feature (motion direction), which equates the effect of directing attention to this stimulus feature; (2) directing attention to a motion direction (fixation vs attend-AP out: directing attention from the fixation condition, where no motion direction is attended, to the pattern moving in the cell's antipreferred direction outside the RF). Here, the RF stimuli are unattended, as the focus of attention remains outside the RF. However, attention is directed from a neutral condition to the antipreferred direction; (3) spatial attention and attending to a motion direction (fixation vs attend-AP in: directing attention from the fixation condition, where no motion direction is attended, to the pattern moving in the cells antipreferred direction inside the RF). Here, attention is directed into the RF and to the neuron's antipreferred direction. These experimental manipulations have proven to be useful at isolating the effects of attending into the RF, or to a motion direction, on neuronal activity in area MT (Treue and Martínez Trujillo, 1999; Martinez-Trujillo and Treue, 2004; Khayat et al., 2010).

Stimuli. The stimuli were back-projected on a screen by a video projector (NEC WT610, 1024 × 768 pixels resolution, 85 Hz). The animals viewed the screen at a distance of 57 cm. The RDPs were generated by plotting bright dots on a dark background with a density of 4 dots per degree^2 within a circular stationary virtual aperture. All dots within one RDP moved coherently at the preferred speed of the recorded neuron (median = 11°/s; range = 4–32°/s), and were replotted at the opposite side when they crossed the border of the aperture. The size of the RDP (1.3–3° diameter) was chosen so that the two patterns fit inside the boundaries of the classical RF excitatory region. Stimulus contrast was measured as the SD of luminance values (Martínez-Trujillo and Treue,

2002; Khayat et al., 2010) and expressed in percentage of the highest value. In the contrast configuration, we used different contrast levels of the test-pattern (0.02, 0.1, 0.3, 0.7, 1.5, 14 and 100% contrast) relative to the contrast level of the AP-pattern. In the direction configuration, we used 7 different motion directions of the test-pattern, from the recorded neuron's preferred direction in steps of 15° until 90° away (Fig. 1*A*) (Khayat et al., 2010).

Recordings and data analysis. Transdural penetrations were made with stainless steel guide tubes (0.3–0.5 mm diameter) through a chamber implanted on top of a craniotomy of the parietal bone that provided access to area MT (Khayat et al., 2010). Spikes and LFPs were recorded simultaneously using standard epoxy-insulated extracellular tungsten electrodes (FHC Inc.; impedance = 1–2 MΩ at 1 kHz), with the guide tube (impedance <0.1 Ω) serving as the reference. We used a Plexon data acquisition system to record and store the neural data (Plexon Instruments). The electrode signal was passed through a headstage with unit gain and then split to separately extract the spike and the LFP components. For the LFPs, the signal was filtered (through hardware filters) between 0.7 and 170 Hz, before being amplified and digitized at 1 kHz. For spike recordings, the signal was filtered between 250 and 8000 Hz, amplified and digitized at 40 kHz. Single-unit spiking activity was then isolated using a window discriminator.

In each recording session, we used an interactive stimulus presentation program and online display of spiking activity to qualitatively assess the recorded neuron's RF location, size, and determine its preferred motion direction and speed. Cells were determined to be from MT according to their response properties (directionality and RF position and size), and to the position of the electrode relative to the superior temporal sulcus assessed through MRI images (Khayat et al., 2010). All recorded neurons had RFs located contralateral to the recording hemisphere/sites and we observed a preponderance of RFs centered between 7° and 12° from the fixation spot in the lower quadrant and horizontal meridian. During the recordings, an infrared eye-tracking device (EyeLink) was used to monitor eye position at a sampling frequency of 200 Hz.

The off-line analysis of the LFP signals was conducted using Matlab software (MathWorks). In each trial, the raw signal was subdivided into several frequency bands (see below) using a second-order, bidirectional, zero phase Butterworth filter, and full-wave rectified. Each rectified bandpass filtered signal was then averaged within two different epochs of 500 ms allowing a spectral resolution of 2 Hz: a baseline period (from −470 to 30 ms relative to stimulus onset), and a response or stimulus presentation time period that started 510 ms after stimulus onset and ended before the occurrence of a stimulus change. The response analysis period therefore started 160 ms after cue-onset, to make sure that the animal had time to direct attention to the target (Khayat et al., 2006; Busse et al., 2008). The rectified bandpass filtered LFP-activity reflects the signal amplitude or strength, and corresponds to the square root of the power on each trial, within a given time period and a given frequency band (Leopold et al., 2003). We determined stimulus-related changes in LFP-power for each frequency band and each trial by computing the base-10 logarithm of the ratio between the stimulus period activity and the baseline period activity [log-ratio = $\log(\text{Power}_{\text{stimulus}}/\text{Power}_{\text{baseline}})$], and then averaging these values across trials and recording sites.

For the stimulus tuning analyses, the signal spectrum between 2 and 170 Hz was first analyzed with a bandwidth of 2 Hz. Then, we recomputed the power within predetermined frequency bands, using different bandwidth settings. For frequencies lower than 30 Hz, we used the conventions of human electroencephalography (EEG bands): δ 2–4 Hz, θ 4–8 Hz, α 8–12 Hz, β 12–30 Hz (Buzsáki, 2006), with the exception that the β band was divided into two sub-bands (β_1 12–20 Hz and β_2 20–30 Hz). For frequencies higher than 30 Hz (γ and above band), we used successive, nonoverlapping 20-Hz-wide bands covering the entire spectrum between 30 and 170 Hz (Liu and Newsome, 2006). For the attentional modulation analyses, in addition to the different EEG bands below 30 Hz, we pooled the higher frequencies (i.e., successive 20-Hz-wide bands) into one broader band signal by recomputing the power between 30 and 120 Hz (γ range or band).

To compare the LFP data with the spiking activity, we also determined in each trial the neuron's averaged firing rate during the baseline and stimulus presentation time periods, and computed the log-ratio. We determined the significance in stimulus-related effects and attentional modulation of LFPs and spiking activity using parametric statistical tests (one-way ANOVA and *t* test). The effects of stimulus configuration and condition on performance were determined using nonparametric statistics (Kruskal–Wallis ANOVA and sign test).

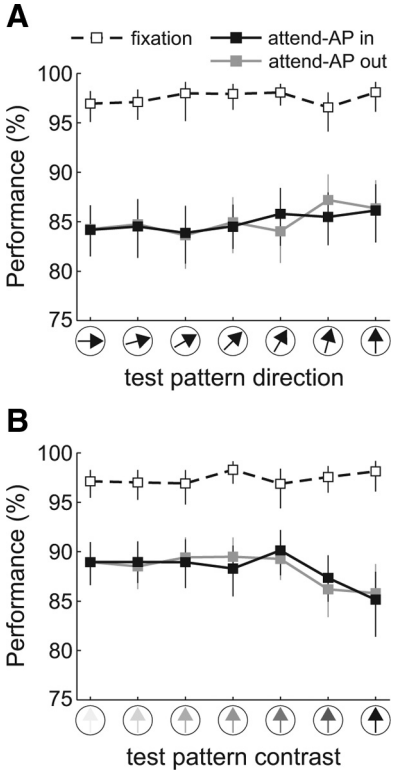


Figure 2. Task performance. ***A***, ***B***, Average performance (hit rate) during the fixation (dashed line), attend-AP in (black), and attend-AP out (gray) conditions as a function of the test-pattern direction (***A***) and contrast (***B***). Error bars denote 95% confidence intervals computed through a bootstrap procedure.

Results

Task performance

We analyzed single-unit firing rate (spiking activity) and LFPs from 81 sites in area MT of two macaques (monkey Se, $n = 48$; monkey Lu, $n = 33$) performing an attention-demanding task. During the task, two pairs of RDPs moving at the neuron's preferred speed were presented, one inside and the other outside the cell's RF. Each pair consisted of a high contrast RDP moving in the neuron's antipreferred direction (AP-pattern) and a test RDP. The latter could have the same contrast as the AP-pattern but moved from trial to trial in different directions (direction configuration), or could move in the neuron's preferred direction but have from trial to trial different contrasts (contrast configuration) (see sketches in Fig. 1*A*). On each trial, the animals had to either attend to the fixation spot and detect a brief change in its luminance (fixation condition, Fig. 1*A*), or to one of the AP-patterns (attend-AP in and attend-AP out condition, Fig. 1*B*) and detect a brief change in its motion direction.

Figure 2 shows the performance (hit rate) in the three behavioral conditions, averaged across sessions ($n = 81$) and monkeys, for the different trial types of the direction (Fig. 2*A*), and contrast (Fig. 2*B*) configurations. The averaged performance for each stimulus pair was similar during both attend-AP conditions ($p > 0.2$, sign test), and lower than during the fixation condition ($p < 10^{-6}$, sign test).

However, in each condition, performance remained relatively constant across levels of each of the manipulated test parameters (direction and contrast) ($p > 0.3$, Kruskal–Wallis nonparametric ANOVA for each condition, with stimulus as a factor), indicating that changes in the direction or contrast of the test-pattern did not significantly influence performance. This allowed us to investigate sensory tuning and attentional modulation of LFPs and spiking activity across different stimulus conditions that yielded similar performance levels (Khayat et al., 2010).

Stimulus tuning of LFP signals

We investigated whether the LFP signal in different frequency bands was tuned for motion direction and contrast. Figure 3*A* shows examples of raw LFP traces and spike trains simultaneously recorded from the same electrode in area MT during trials with two high-contrast RDPs moving in opposite directions (preferred and antipreferred) presented inside the unit's RF (see top panel).

We first examined the shape of the LFP power spectrum from 2 to 170 Hz by bandpass filtering the signal with a 2 Hz frequency resolution (see Materials and Methods). On each trial, the power was determined separately for each 2 Hz band during the 500 ms baseline period before stimulus onset, and during the period from 510 to 1010 ms after stimulus onset (Fig. 3*A*, shaded areas), and then averaged across trials. Figure 3, *B* and *C*, shows the LFP signal power (mean ± SEM) during fixation trials of the same example site depicted in Figure 3*A*, and averaged across sites corresponding to each monkey. In all cases, the power was dominated by low-frequency components, decreasing at higher frequencies during both the baseline (gray), and the stimulus period (black).

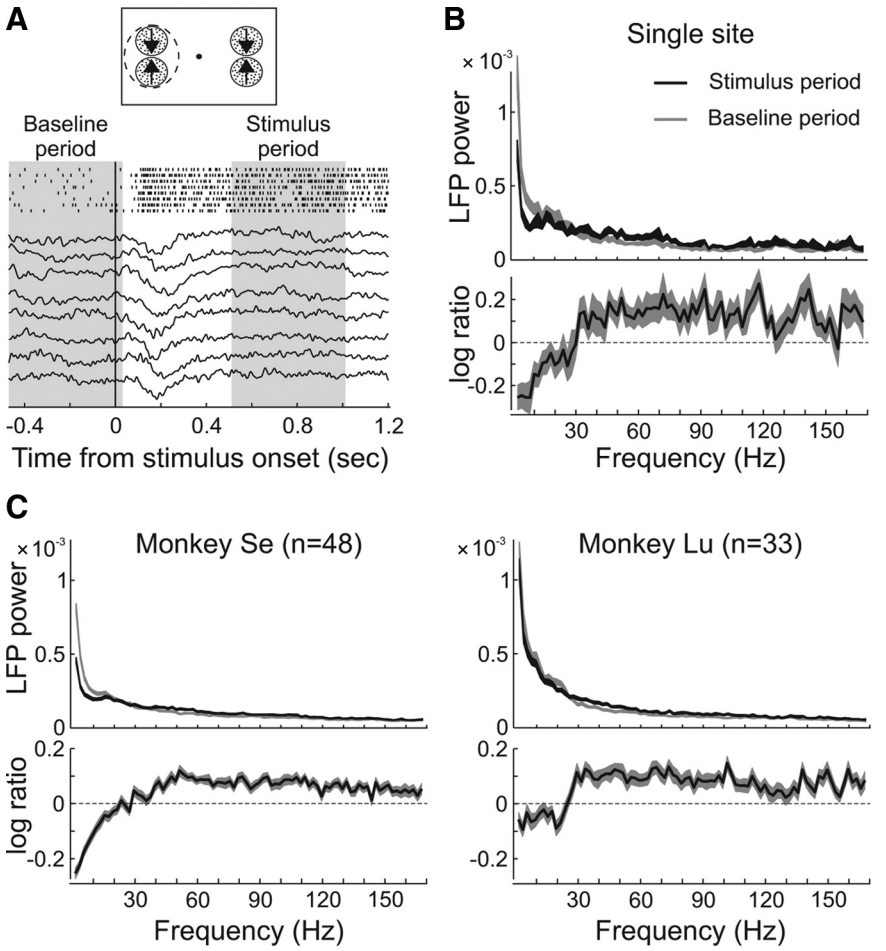


Figure 3. LFP analysis. ***A***, LFP activity traces of single fixation trials at an example-recording site. The spike trains of the simultaneously recorded single unit are also shown. The recorded activity is aligned to the onset of the stimulus depicted on top. The shaded areas show the baseline and the stimulus analyses periods. ***B***, Mean LFP-power across frequency (2 Hz bandwidth) during the baseline (gray) and the stimulus response (black) periods, averaged across fixation trials corresponding to the example-recording site. The thickness of the lines depicts ±1 SEM. The bottom shows the relative-power during stimulus presentation, computed as the base-10 logarithm of the ratio between the power in the two analysis periods. ***C***, LFP-power averaged across the recording sites of monkey Se ($n = 48$) and Lu ($n = 33$) in the two time periods during fixation trials. The relative power (log-ratio, bottom) was computed for each recording site (as in ***B***) and then averaged across sites.

To determine stimulus-related changes in LFP-power, we computed in each trial the base-10 logarithm of the ratio between the power during both time periods [$\log_{10}(\text{Power}_{\text{stimulus}}/\text{Power}_{\text{baseline}})$], and then averaged these values across trials and recording sites (Fig. 3*B*,*C*, bottom). Observe that this analysis shows the average log-ratio across trials and not the log-ratio of the mean values plotted in the top. The results of the single site example, as well as those averaged across the population of sites of each individual monkey (monkey Se, $n = 48$; monkey Lu, $n = 33$) followed the same profile: stimulus presentation reduced the power at frequencies below ~30 Hz, and increased it at higher frequencies.

We conducted the same analysis for the different stimulus pairs of the direction and contrast configuration (Fig. 4*A*,*D*) by pooling the data from both animals ($n = 81$). Figure 4 illustrates the mean relative power (i.e., log-ratio) across frequencies during fixation trials with different directions (Fig. 4*A*) and contrasts (Fig. 4*D*) of the test-pattern. In general, stimulus presentation reduced the power at frequencies below ~30 Hz (blue colors), and increased it at higher frequencies (red and yellow). Moreover, changing the test-pattern direction away from the preferred direction or lowering its contrast progressively decreased the magnitude of the relative LFP-power at frequencies above ~30 Hz (γ range) (Fig. 4*A*,*D*, red-yellow gradient from right to left), but had not effect at frequencies below ~30 Hz. The results were similar in both animals.

To quantify these observations, we computed for each site the relative LFP-power in predetermined frequency bands. For frequencies above 30 Hz, we used successive 20-Hz-wide frequency bands (30–50, 50–70, 70–90, 90–110, 110–130, 130–150, 150–170 Hz) (Liu and Newsome, 2006). For frequencies below 30 Hz, we used the classical EEG bands: 2–4 Hz (δ), 4–8 Hz (θ), 8–12 Hz (α), 12–20 Hz (β_1), and 20–30 Hz (β_2) (Buzsáki, 2006). Across the population of sites the relative LFP-power in the classical EEG bands did not significantly change with changes in the test-pattern direction (Fig. 4*B*) or contrast (Fig. 4*E*) ($p > 0.5$, one-way ANOVA). On the other hand, in the above-30 Hz bands, the relative power across sites decreased as the test-pattern's motion direction deviated from preferred (Fig. 4*C*), or decreased contrast (Fig. 4*F*) ($p < 0.0005$, one-way ANOVA).

We further measured stimulus selectivity of the LFP signal by computing for each recording site and frequency band the slope

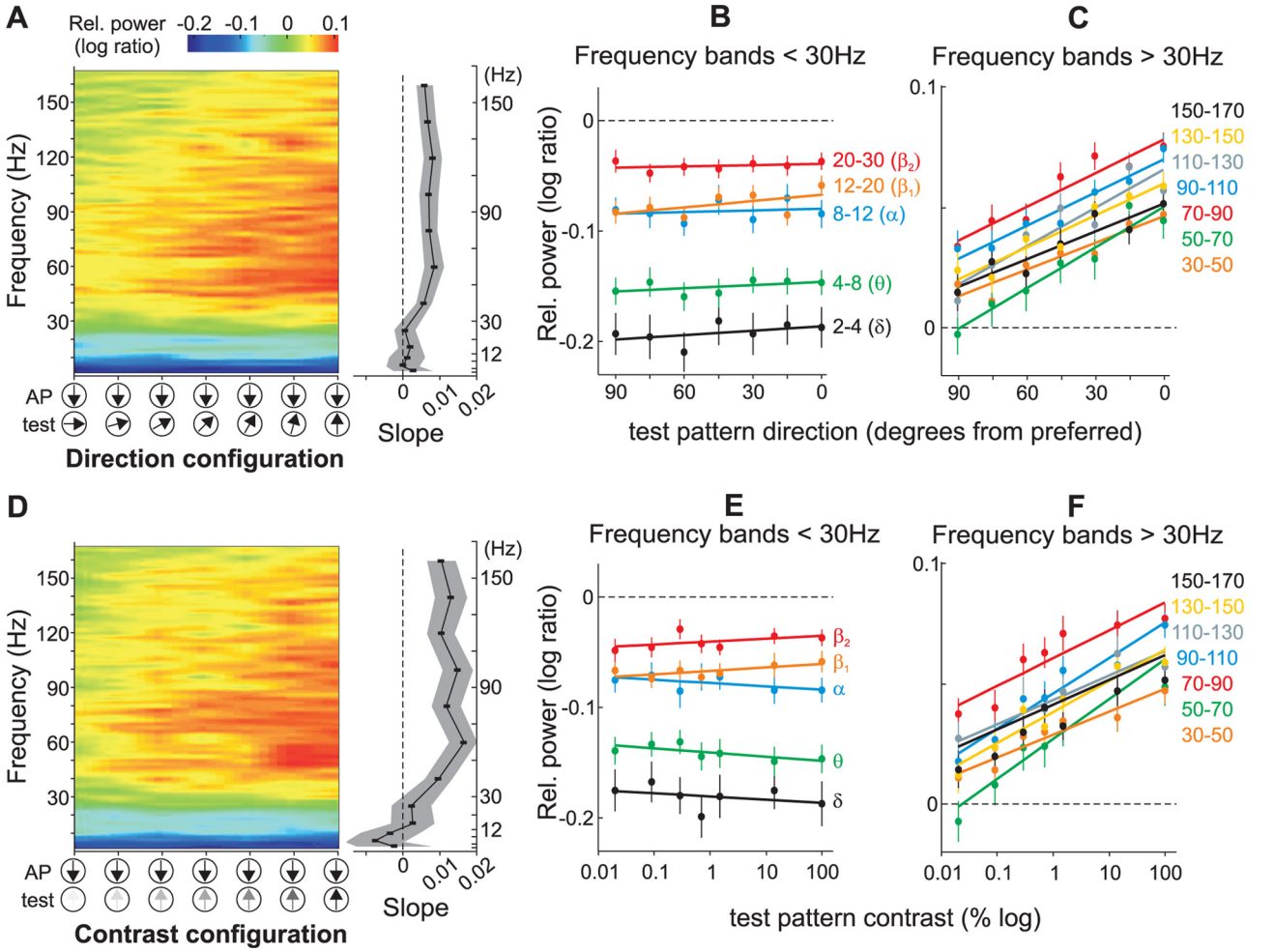


Figure 4. LFP tuning in the direction and contrast configuration. ***A***, Color plot showing the relative LFP-power across frequency (2 Hz bandwidth) and test-pattern motion directions during fixation trials averaged across recording sites ($n = 81$). The color represents the change in power during stimulus presentation relative to the baseline activity (log-ratio). For visual display purposes, the data points across frequency and stimulus pairs were interpolated. The graph on the right shows the mean slope (95% confidence interval shown in gray) of a regression line fitted to the data in predetermined frequency bands (see ***B*** and ***C***, and Results). The slope was computed for each recording site and then averaged across sites. ***B***, ***C***, Mean relative LFP-power in different frequency bands below (***B***) and above (***C***) 30 Hz, for the population of recording sites ($n = 81$) as a function of the test-pattern motion direction. Error bars, SEM. Note the different scale between the panels. ***D–F***, Data in the contrast configuration. Same conventions as above.

of a regression line fitted to the data corresponding to the direction and contrast configurations (see lines in Fig. 4*B*,*C*,*E*,*F*). In both configurations, the average slope across recording sites in the below-30 Hz bands was not different from zero ($p > 0.05$, *t* test; r^2 range: 0.04–0.5), indicating that the signal power was not tuned for direction and contrast. On the other hand, for the above-30 Hz bands the average slopes were significantly larger than zero ($p < 10^{-6}$, *t* test; r^2 range: 0.8–0.94), indicating that the relative power was tuned for both direction and contrast (Fig. 4*A*,*D*, see right panels).

This LFP-signal tuning in the above-30 Hz frequencies parallels the tuning found in the neuronal spiking activity recorded from the same electrode (Fig. 5*A*). To facilitate the comparison with the LFP data, we determined the neurons relative firing rate by computing the base-10 logarithm of the ratio between the spiking activity during the stimulus and baseline periods in each trial, and then pooling across trials and cells. We then computed the correlation between spikes and LFPs by plotting the relative LFP-power in each frequency band as a function of the relative firing rate, and fitting a regression line to the data. In the direction (Fig. 5*B*, left) and contrast (Fig. 5*B*, right) configurations, the firing rate was significantly and positively correlated with the LFP-power in the above-30 Hz frequency bands (r^2 range: 0.78–0.94; $p < 0.005$, *t* test), but not in the below-30 Hz bands (r^2, range 0.04–0.3; $p > 0.05$, *t* test) (Fig. 5*C*, see mean linear regression slopes ± 95% confidence interval, for each band).

Attentional modulation of low-frequency (<30 Hz) LFP signals

To examine whether visual attention modulated the LFP signal in our sample of recording sites, we compared the relative LFP-power in the different bands between the three behavioral conditions (fixation, attend-AP in and attend-AP out). Figure 6 shows the relative LFP-power in the δ (2–4 Hz), θ (4–8 Hz), α (8–12 Hz), β_1 (12–20 Hz), and β_2 (20–30 Hz) frequency bands averaged across recording sites ($n = 81$) in the three behavioral conditions of the direction and contrast configuration, as well as attentional effects. In none of these lower frequency bands did the magnitude of the LFP-power vary significantly with changes in the test-pattern direction (Fig. 6*A*, left), or contrast (Fig. 6*B*, left) ($p > 0.5$, one-way ANOVA) when the animals had to attend to the fixation spot (black dashed line, fixation), or directed attention to the AP-pattern inside (black solid line, attend-AP in) or outside (gray solid line, attend-AP out) the RF. However, except

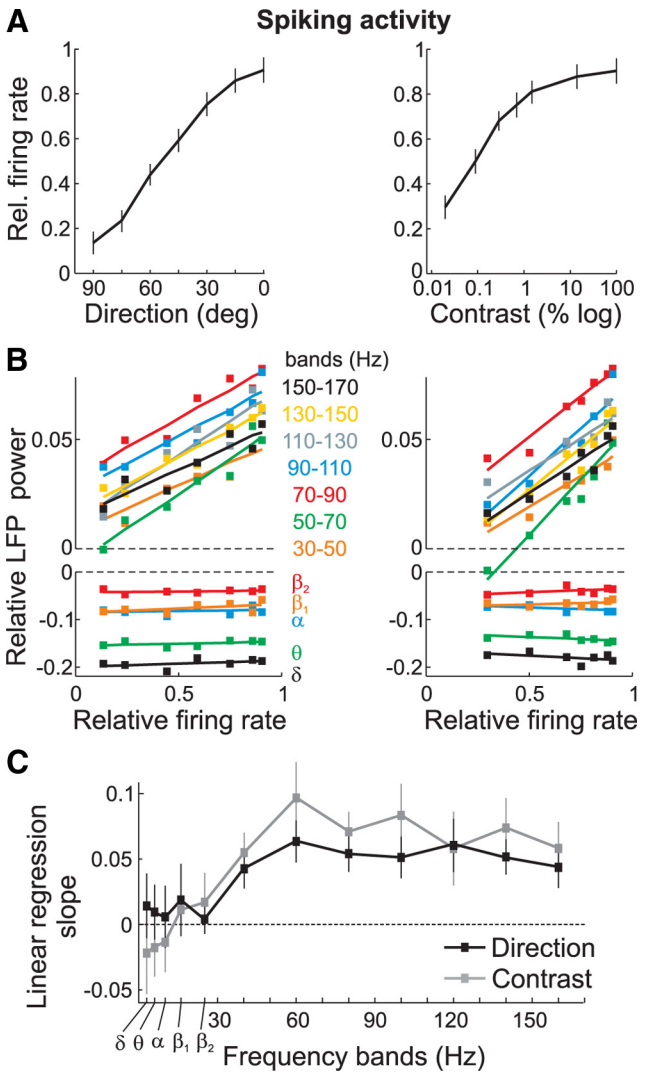


Figure 5. Relationship between LFPs and spiking activity. ***A***, Relative firing rate during fixation trials averaged across recording sites ($n = 81$) in the direction (left) and contrast (right) configuration. Error bars, SEM. ***B***, Relative LFP-power in the different bands versus the relative firing rate, in the direction (left) and contrast (right) configuration. A regression line was fitted to the data points. ***C***, Slope of the linear regression fits shown in ***B*** for the different bands, in the direction (black) and contrast (gray) configuration. Error bars, 95% confidence interval.

for the θ band, the strength of the LFP-power across all test-pattern directions and contrasts appears to be reduced when the animals directed attention inside the RF (attend-AP in) relative to the other conditions.

We quantified these observations by computing, for each recording site and stimulus combination, the difference in LFP-power between the attend-AP in and attend-AP out or fixation condition, and converting these values to percentages of the LFP-power in the attend-AP out or fixation condition, depending on the comparison to be made (100 × [attend-AP in − attend-AP out]/attend-AP out), or (100 × [attend-AP in − fixation]/fixation). In the δ, α, β_1, and β_2 bands, but not in the θ band, the difference in the LFP-power pooled across sites (Fig. 6, colored panels) was significantly lower than zero when comparing the attend-AP in and attend-AP out condition (blue), and the attend-AP in and fixation condition (red), for the majority of stimulus pairs (open symbols, $p < 0.05$, t test). The magnitude of this effect extended from ~20% (in the δ band) to ~75% (in the β_2 band) suppression in attend-AP in trials compared with attend-AP out or fixation trials. This cannot be explained by variations in the animals' behavior, since the task in both attend-AP in and out trials, as well as the animal performance were similar (Fig. 2). Rather, it appears to be caused by directing attention into the RF, because the suppression was equally strong when comparing attend-AP in to fixation, and attend-AP in to attend-AP out, while there was not effect when comparing attend-AP out to fixation. We will further investigate this issue in a separate section (see end of the Results).

Attentional modulation of LFP signals in the γ band: direction configuration

To examine attentional effects in frequency bands above 30 Hz, we determined the relative power in the γ band (30–120 Hz) in the three behavioral conditions. We pooled the data across these frequencies since in our previous analysis the LFP-power was similarly tuned for motion direction and contrast in all bands from 30 to 120 Hz. Additionally, previous studies of attention have described a signal modulation in this frequency range (Fries et al., 2001, 2008). In the three conditions, the relative power decreased with changes in the test-pattern's direction away from preferred (Fig. 7*A*, $p < 10^{-5}$, one-way ANOVA). This was anticipated from the results shown in Figure 4, and reflects the selectivity of the γ-band signal power for motion direction (Liu and Newsome, 1996). Relative to fixation trials, we found that in both attend-AP in and out conditions the power was suppressed for directions of the test-pattern similar to the preferred (rightmost data points) but enhanced for directions away from the preferred (leftmost data points).

This effect is better illustrated in Figure 7*A* (right), which shows the average difference in signal power for each comparison and test-pattern direction. The average difference between the attend-AP in and fixation condition (red), and between the attend-AP out and fixation condition (green), shifted from negative to positive values as the direction of the test-pattern deviated from the preferred direction ($p < 0.05$, one-way ANOVA). Note that the magnitude of these effects is relatively weak compared with that observed in the lower frequency bands. However, relative to the percentage of signal modulation due to changes in the stimulus direction in the fixation condition (i.e., difference between the signal evoked by the AP + preferred test-pattern and AP + 90° away-from-preferred test-pattern), the γ power in the attend-AP in (or attend-AP out) condition was modulated within a range of 39% (26%) relative to fixation trials. The modulation shifted from 19% (10%) suppression (rightmost data point, Fig. 7*A*, right panel) to 20% (16%) enhancement (leftmost data point).

We have previously suggested that this direction-dependent effect is due to the effects of attention on input signals into area MT (Khayat et al., 2010). Basically, when direction-selective inputs activated by the test stimulus maximally differ in their preferred direction from the attended direction (180° apart, i.e., preferred direction test-pattern vs AP-pattern), they would therefore be maximally suppressed relative to when they become more similar (90° apart). This would result in the test-pattern direction-dependent modulation shown in Figure 7. If one assumes that LFPs in the γ band represent, to a certain extent, the strength of inputs into a local circuit, and therefore into neurons within that circuit, our results would agree with this hypothesis, and with recently proposed computational and theoretical models in which the strength of feature-selective inputs into a neuron is modulated by attention (Reynolds and Heeger, 2009) (see Discussion).

When controlling for the effects of attending to a motion direction by comparing attend-AP in to attend-AP out trials

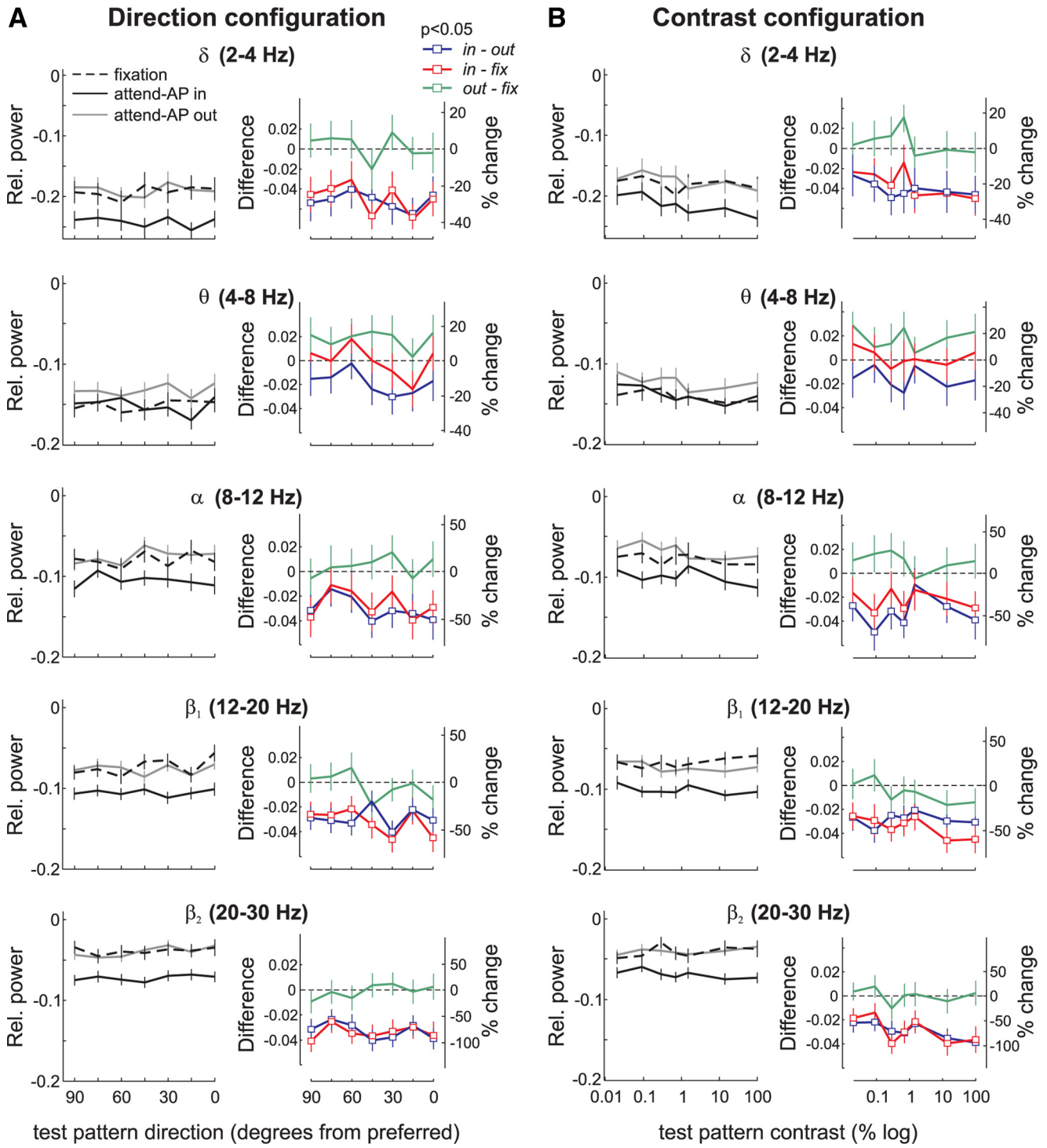


Figure 6. Attentional modulation of LFPs in the low-frequency bands. ***A***, Direction configuration data in δ, θ, α, β_1 and β_2 frequency bands. The left panels show the mean relative power ($n = 81$) in the three behavioral conditions for the different test-pattern directions. Black, Attend-AP in; gray, attend-AP out; dashed line, fixation. The right panels show the magnitude of attentional effects expressed as the difference (left axis) and the percentage change (right axis) in power for the comparison attend-AP in versus attend-AP out (in-out, blue), attend-AP in versus fixation (in-fix, red), and attend-AP out versus fixation (out-fix, green). Square data points denote significant differences in power between conditions ($p < 0.05$, *t* test). Error bars, SEM. ***B***, Contrast configuration data in δ, θ, α, β_1 and β_2 frequency bands. Same conventions as in ***A***.

(spatial attention), the shift in modulation disappeared, and there was no significant interaction between test-pattern direction and the modulation in the γ band (blue, $p > 0.8$, one-way ANOVA, Fig. 7*A*). The difference in γ power pooled across all stimulus combinations was significantly lower than zero (-0.002 ± 0.001, mean $\pm$ SEM, $p < 0.05$, *t* test), and represented ~6% mean suppression relative to the magnitude of the tuning, which indicates that the γ-band LFP-power is generally suppressed when directing attention to the AP-pattern located inside the recorded neuron's RF. These effects of spatial attention, however, did not reach significance when analyzing each stimulus combination separately ($p > 0.05$, *t* test), and were

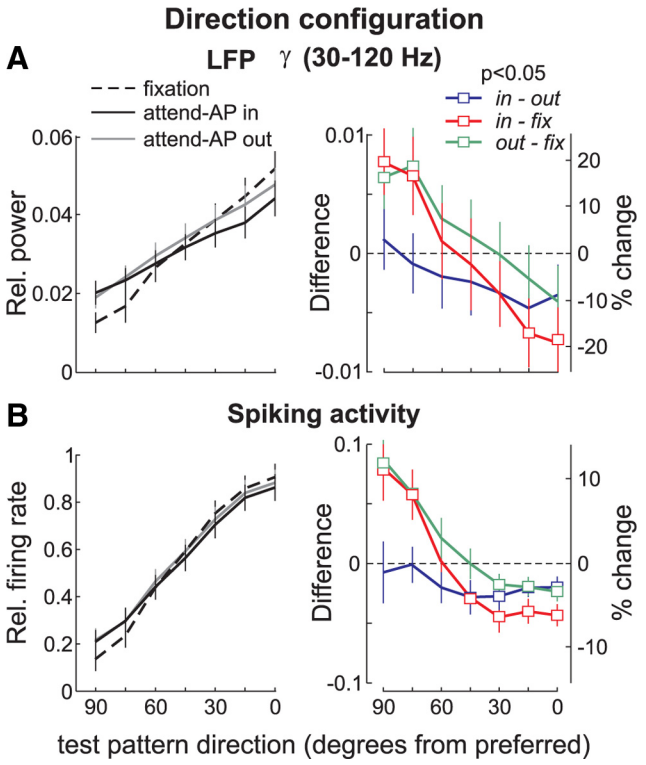


Figure 7. Attentional modulation of LFPs in the γ band (30–120 Hz) and spiking activity, in the direction configuration. The left panels show the relative γ power (***A***) and the relative firing rate (***B***) across the population of recording sites ($n = 81$), in the three behavioral conditions. Black, Attend-AP in; gray, attend-AP out; dashed line, fixation. The right panels show the difference (left axis) and the percentage change (right axis) in γ power (***A***) and in firing rate (***B***), for the various comparisons. Attend-AP in versus attend-AP out (blue). Attend-AP in versus fixation (red). Attend-AP out versus fixation (green). Squares denote data points with significant differences in power between conditions ($p < 0.05$, *t* test). Error bars, SEM.

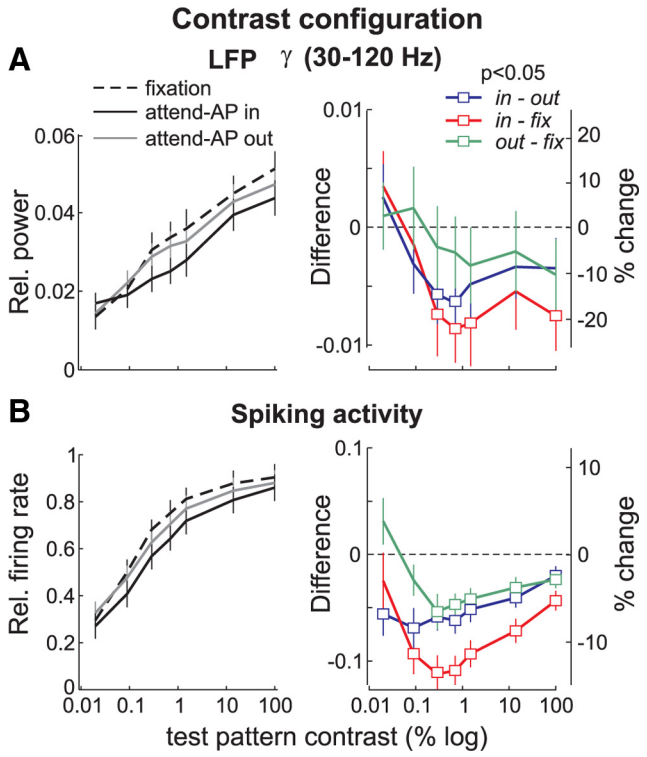


Figure 8. Attentional modulation of LFPs in the γ band (30–120 Hz) and spiking activity, in the contrast configuration. Same conventions as in Figure 7.

considerably weaker than those in the corresponding low-frequency power (Fig. 6*A*).

Comparison with the spiking activity

The firing rate modulation between attend-AP in (or attend-AP out), and the fixation condition also shifted from suppression to enhancement as the direction of the test-pattern deviated from preferred (Fig. 7*B*, red and green, right) ($p < 10^{-6}$, one-way ANOVA). The suppression/enhancement pattern resulted in a total modulation of 16% (in-fix, red) and 14% (out-fix, green) change relative to the strength of the tuning. When isolating the effects of spatial attention (attend-AP in vs attend-AP out), the firing rate was generally suppressed when attention was directed to the AP-pattern inside the RF. We found that the difference (blue) was significantly below zero for stimulus pairs in which the test-pattern's direction was more dissimilar to the attended direction (open symbols; $p < 0.05$, *t* test). For the other test-pattern directions, the suppression was considerably attenuated ($p > 0.05$, *t* test). There was no significant difference in the modulation as a function of the test-pattern motion direction ($p > 0.5$, one-way ANOVA). A more extensive report of these findings can be found in the work of Khayat et al. (2010). Our interest here is to compare these modulations with the ones of the LFP-signal power shown in Figure 7*A*.

In summary, the LFP-power modulation in the γ band, but not in the other frequency bands, followed the modulation of the neurons spiking activity (Figs. 6*A*, 7). In terms of relative strength, the LFP modulation in all frequencies (but the θ band) was stronger than the modulation of the neurons spiking activity. However, one must interpret this later finding cautiously, mainly because the quantitative details of the way in which LFP signals relate to spiking activity are, so far, poorly understood.

Attentional modulation of LFP signals in the γ band: contrast configuration

In the contrast configuration, we found that the relative power in the γ band (30–120 Hz) was generally suppressed in the attend-AP in compared with the other two conditions (solid black line, Fig. 8*A*, left). Moreover, we found that the strength of attentional modulation varied with contrast (red and blue, Fig. 8*A*, right); the differences between conditions were the largest for stimuli of intermediate contrast, representing 22% (in-fix) and 16% (in-out) change in relative power with respect to the strength of contrast tuning. Note that when comparing the attend-AP out and fixation conditions attentional effects are considerably attenuated, loosing significance (green, $p > 0.05$, *t* test for each stimulus pair). Importantly, neither for the contrast configuration, nor for the direction configuration (Fig. 7), could these attentional effects reflect changes in the animals' performance since hit rate across all attend-AP in and attend-AP out trials was almost identical, while in fixation trials it was significantly higher (Fig. 2).

Comparison with the spiking activity

In the contrast configuration, spiking activity was suppressed in both attend-AP in and out conditions, relative to fixation (Fig. 8*B*). This suppression was significant across most contrast levels (right panel, open symbols, $p < 0.01$, *t* test). Moreover, this effect was strongest for intermediate contrast stimuli, particularly when comparing the fixation condition to either attended condition

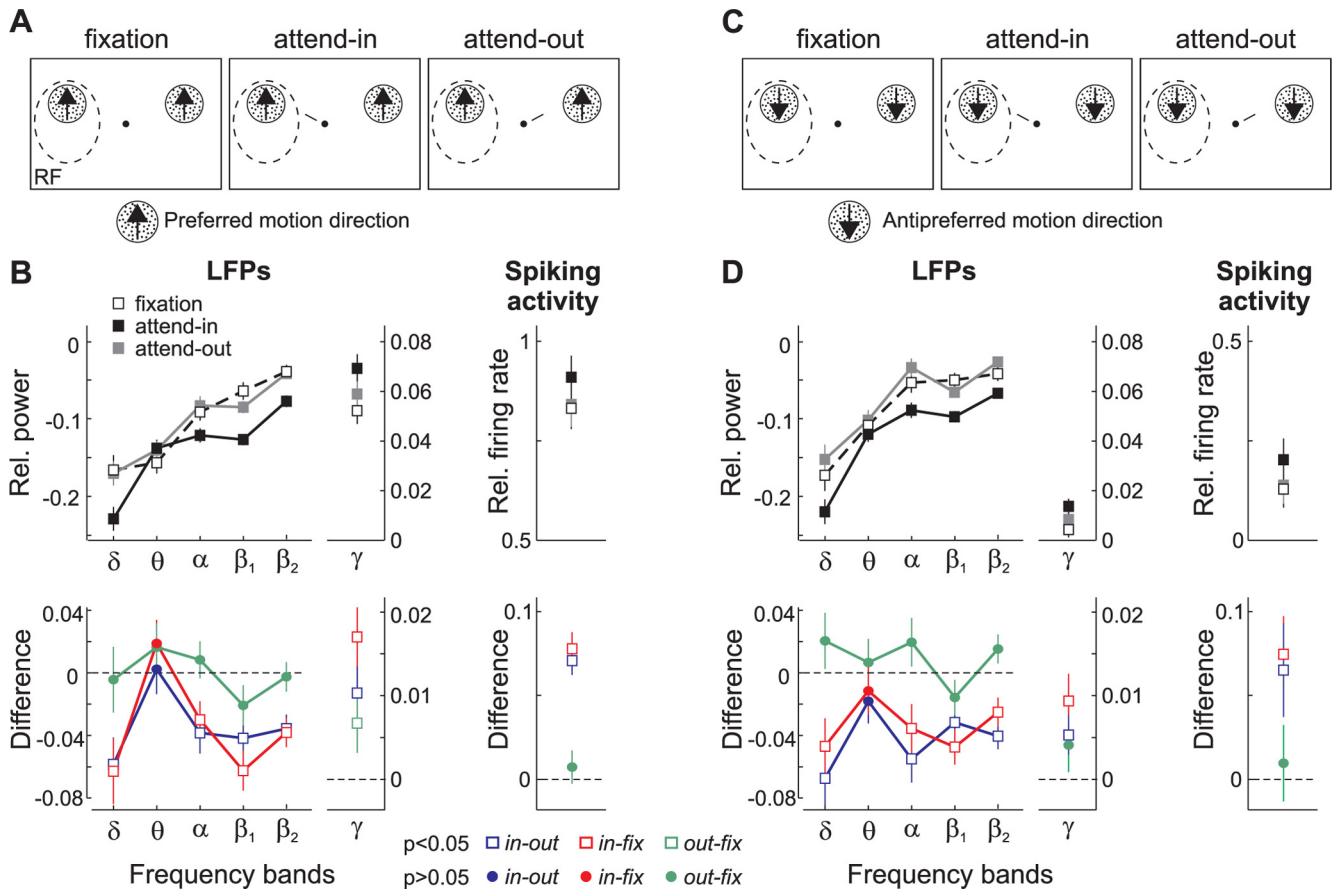


Figure 9. Attentional modulation of population activity ($n = 81$) evoked by either the preferred or antipreferred motion direction. ***A***, Stimulus display showing one pattern moving in the preferred direction located inside the RF (dashed circle), and another identical one located in the other hemifield. The animal ignored the patterns (fixation condition) or directed attention to either one of them (attend-in and attend-out conditions). ***B***, The top shows the relative power in the δ, θ, α, β_1, β_2 and γ bands (left), and relative firing rate (right) during the three conditions with the preferred pattern alone. Black squares, Attend-in; gray squares, attend-out; white squares, fixation. The bottom shows the difference between conditions for each measurement of cortical activity. Blue, Attend-in versus attend-out; red, attend-in versus fixation; green, attend-out versus fixation; open squares denote significant differences in power between conditions ($p < 0.05$, *t* test). Circles denote nonsignificant ones ($p > 0.05$). Error bars, SEM. ***C***, Stimulus display with one pattern inside the RF moving in the antipreferred direction. ***D***, Relative power in the different bands and relative firing rate during the three conditions with the antipreferred pattern alone (top), and the corresponding differences in activity between conditions (bottom). Same conventions as in ***B***.

(red and green, $p < 0.0005$, one-way ANOVA). Here, the amount of change in the firing rate reached 14% (in-fix, red) and 8% (out-fix, green) at intermediate contrast (third data point from the left).

In general, the shape of the modulation of the spiking activity and of the LFP-power in the γ band was similar. Although the magnitude of these effects was typically larger in the LFPs relative to the spikes (as it was for the direction configuration), the statistical significance was reached more often when comparing the magnitude of the spiking activity between the different conditions. This may be due to a larger variability in the LFP-power values across individual sites compared with the spikes (see error bars). In the lower frequency bands, the LFP-power modulation did not correlate with the modulation of the spiking activity.

Attentional modulation of LFP-power and spiking activity with one pattern in the RF

So far, our results show that attending to the AP-pattern could produce either an enhancement or a suppression of both the γ-band LFP-power, and the neurons firing rate. On the other hand, the power in the δ, α, β_1, and β_2 band was suppressed when attention was directed to the AP-pattern inside the RF, and was not influenced by the different test-pattern directions or contrast.

To further test whether the suppression of power in these low-frequency bands was due to directing attention into the RF, we used an additional stimulus configuration that differed from the previous ones. During the recording sessions, we included trials in which we presented a single RDP inside the RF and a second identical one in the opposite hemifield (Treue and Martínez Trujillo, 1999). Both patterns moved either in the preferred (Fig. 9*A*) or antipreferred (AP) direction (Fig. 9*C*) of the recorded neuron. These trials were randomly interleaved with the trials of the other configurations. The animals were instructed to direct attention either to the pattern located inside (attend-in), or outside (attend-out) the RF, or to the fixation spot (fixation).

Figure 9*B* shows the LFP-power in the different frequency bands and the spiking activity averaged across recording sites ($n = 81$) during the presentation of the preferred motion direction. When attention was directed inside the recorded neurons' RFs the relative LFP-power (black data points) in the δ, α, β_1, and β_2 bands, but not in the θ band, was reduced compared with the fixation (white) and attend-out (gray) condition. The bottom shows the differences between conditions, which were significantly lower than zero (square symbols, $p < 0.05$, *t* test) in all bands below 30 Hz, except in the θ band, when comparing attend-in and attend-out trials (blue), as well attend-in and fixation trials (red). On the other hand, the power between attend-out and fixation trials did not differ (out-fix, green and circle symbols, $p > 0.05$, *t* test).

When the antipreferred pattern was presented in isolation inside the neuron's RF, the LFP-power in these low-frequency bands was modulated in a similar manner (Fig. 9*D*). These results parallel those observed when two patterns were located inside the neurons' RFs (Fig. 6), and indicate that directing attention into the RF reduces δ, α, β_1, and β_2 LFP activity in the neuron's vicinity, regardless of the stimulus feature or number of stimuli.

On the other hand, both the LFP-power in the γ band and the spiking activity were enhanced during the attend-in condition (black) relative to the other two conditions, when either the preferred (Fig. 9*B*) or the antipreferred pattern (Fig. 9*D*) was located in isolation inside the RF. Note that the γ power and spiking activity during trials with direction and contrast stimulus pairs (Figs. 7, 8) fall between the level of activity evoked by the preferred pattern alone (Fig. 9*B*), and the antipreferred pattern alone (Fig. 9*D*). This shows that in the previous configurations, adding a second test-pattern (with a given direction and contrast) to the AP-pattern inside the RF increased the γ power and firing rate, and also reflects the tuning of both measurements for direction and contrast.

When either pattern was presented in isolation inside the RF, the average difference in the γ power and spiking activity between attend-in and attend-out or fixation trials (blue and red) was significantly higher than zero (see bottom, square symbols, $p < 0.05$, *t* test), reflecting an enhancement in both measurements when attention was directed inside the RF (Treue and Martínez Trujillo, 1999). We also isolated the effects of directing attention to a motion direction, while keeping the focus of attention outside the RF (attend-out vs fixation, green data points). We found that the γ-power and spiking activity evoked by the unattended pattern (preferred or antipreferred) inside the RF was generally enhanced when attention was directed to a similar pattern located outside the RF relative to the fixation condition. However, this direction-dependent effect was relatively weak and reached significance only in the γ-power activity evoked by the preferred pattern (Fig. 9*B*, green square, $p < 0.05$, *t* test).

In summary, these results show that directing spatial attention to a single stimulus inside the RF generally enhanced the neurons' spiking activity and the strength of LFP-power in the γ band, while consistently decreased the power in δ, α, β_1, and β_2 bands. We did not find a significant influence of attention on the LFP-power in the θ band.

Discussion

Our study demonstrated that: (1) LFP-power in area MT was, in the γ-frequency range, but not in the δ, θ, α, β_1, and β_2 bands, tuned for the motion direction and contrast of moving RDPs, similarly to the neurons firing rate, (2) directing attention into a neuron's RF decreased LFP-power in the bands below 30 Hz (except the θ band), and could either increase/decrease power in the γ band, (3) directing attention to a motion direction outside the RF had no effect in LFP-power below 30 Hz, but could either increase or decrease γ depending on the relationship between the attended direction, and the direction of the stimuli located inside the neurons' RF, d) in the γ band, the effects of attention closely matched those observed in the neurons firing rate.

Stimulus tuning of LFPs

A previous study in area MT has reported tuning of the LFP-power in the γ and above frequencies for the direction and speed of moving RDPs (Liu and Newsome, 2006). Our findings agree with this report and also show that the γ-band power was tuned for stimulus contrast, as previously reported in area V1 (Henrie and Shapley, 2005). Other studies have described stimulus selectivity in this band for orientation, and spatial/temporal frequency in area V1 (Frien et al., 2000; Kayser and König, 2004), and for objects in inferior temporal (IT) cortex (Kreiman et al., 2006). Thus, gamma-band power selectivity for stimulus features is commonly found in visual cortical areas (for review, see Berens et al., 2008).

In our study, the LFP-power in the δ, θ, α, β_1, and β_2 bands remained invariant to the different RDPs direction and contrast, which is consistent with previous studies in areas MT (Liu and Newsome, 2006), and V1 (Henrie and Shapley, 2005). However, it apparently differs from a recent report in area V1 showing stimulus tuning in the low frequencies (Katzner et al., 2009). We believe that this discrepancy may be due to the fact that Katzner et al. (2009) used stimuli evoking strong transient neuronal responses, while we purposely avoided using these stimuli and concentrated in analyzing sustained responses, long after stimulus onset.

In accordance with our findings, Belitski et al. (2008) found that area V1 LFP oscillations in frequencies <40 Hz show very little signal and noise correlation with frequencies >40 Hz. They suggested that low-frequency LFPs reflect neural processes that in natural conditions are fully decoupled from those giving rise to spikes and to γ LFPs. They further hypothesized that γ-band LFPs and spikes are generated within the same network. Supporting this idea, a recent study in macaque IT cortex demonstrated that sensory adaptation similarly affects spiking activity and γ-band LFPs (De Baene and Vogels, 2009). Our results are fully compatible with this hypothesis.

Effects of attention in the γ band

In the direction configuration of the first experiment, shifting attention from the fixation spot to the AP-pattern suppressed both the γ band and the spiking activity when the test-pattern moved in the preferred and similar directions. When the test-pattern moved in directions orthogonal to the preferred, and therefore closer to the antipreferred, both measurements were enhanced. This result was similar (albeit different in magnitude) regardless of whether the attended AP-pattern was inside, or outside the neurons' RF. We hypothesize that this modulation results from a direction-dependent attentional mechanism that enhances the strength of inputs selective for directions similar to the attended antipreferred direction, while suppresses the strength of inputs selective for more dissimilar directions (Khayat et al., 2010). This agrees with recently proposed models, in which attention modulates the strength of sensory inputs into visual neurons, and consequently the amount of response normalization the cells undergo (Reynolds and Heeger, 2009).

It is reasonable to assume that most direction-selective inputs into MT arise from populations of direction-selective neurons in upstream areas, e.g., V1, V2, or V3 (Born and Bradley, 2005). Since γ-band LFPs are tuned for direction, one may hypothesize that they represent the contribution of these direction-selective inputs into an MT recording site (Khawaja et al., 2009). The effects of attention in the γ-band may arise from increases or decreases in the synchronized activity of the corresponding upstream neural populations with RFs at attended locations (Frien et al., 1994; Fries, 2009). Such a mechanism has been proposed as underlying increases in the synchronous firing of V4 neurons with attention (Fries et al., 2001, 2008; Taylor et al., 2005). Other potential explanations are changes in the synchronous behavior of neurons within an MT directional column via interactions with neighboring cells, or with neurons in downstream areas

through feedback projections (for review, see Tiesinga and Sejnowski, 2009; Gregoriou et al., 2009).

Effects of attention in the low-frequency bands

One ubiquitous finding across all stimulus configurations used in this study was that directing attention into the neurons' RF suppressed LFP-power in the δ, α, β_1, and β_2 bands, but not in the θ band. This effect was independent of the stimulus configuration (one- or two-stimuli in the RF), and disappeared when the animals switched attention between the fixation spot and the pattern outside the RF (attending to a direction). Importantly, our experimental manipulations led to situations in which spatial attention either decreased (contrast configuration), or increased (single pattern inside the RF configuration) the firing rate of an MT neuron when directed into its RF. In all these situations, however, LFP-power in the δ, α, β_1, and β_2 bands was suppressed, suggesting that this effect was not linked to increases or decreases in firing rate. Moreover, the magnitude of the modulation in these bands was stronger than in the γ band.

A study in macaque area V4 using a "one stimulus inside and one outside the RF" configuration (Fries et al., 2008), as well as EEG/MEG studies in humans (Worden et al., 2000; Kelly et al., 2006, 2009; Thut et al., 2006; Siegel et al., 2008), reported similar results. However, Lakatos et al. (2008) reported in area V1 that attention could also increase the power of low-frequency oscillations in tasks where the occurrence of a stimulus within a rhythmic stream is highly predictable. They suggested that attention could increase or decrease δ power in such a manner that a phase of high excitability would coincide with the stimulus occurrence; i.e., enforcing oscillatory entrainment to a task-relevant input stream. On the other hand, when the behaviorally relevant event is unpredictable in time—as was the direction change of our experiment—attention operates best in a "continuous mode" characterized by extended increase in γ and suppression of low-frequency oscillations (Schroeder and Lakatos, 2009). The results of Lakatos et al. (2008), together with the ones of Katzner et al. (2009), suggest that LFP-power in the low frequencies can also vary according to the nature of the sensory input and the task (see also Buschman and Miller, 2009).

Interestingly, despite the observed effects of attending to a stimulus direction outside the RF on firing rate and γ LFPs, we did not find any correlate of such effects in the low frequencies. A previous study in area V4 (Bichot et al., 2005) reported clear enhancements in firing rate and synchrony between spikes and γ-band activity when attending to a stimulus feature, but no effect in the low frequencies. These results suggest that, unlike spatial attention, feature-dependent attentional mechanisms may not "target" low-frequency oscillations.

One possible explanation for the latter phenomenon is linked to the proposal that low-frequency LFPs reflect signal pooling (through lateral connections) across neurons within a cortical region when feedforward inputs are weak, favoring the detection of low intensity visual signals (Nauhaus et al., 2009). If spatial attention increases the strength of feedforward (upstream) inputs to neurons with RFs at the attended location (Reynolds et al., 2000; Martínez-Trujillo and Treue, 2002; Reynolds and Heeger, 2009; Khayat et al., 2010), it will decrease the need for signal pooling. Because inputs' strength and low-frequency power are inversely correlated, one would expect a decrease in low-frequency oscillations in the region targeted by the strengthened inputs. On the other hand, because more global feature-dependent mechanisms produce, within the same cortical region, increases and decreases in inputs' strength, depending on the selectivity of the inputs (i.e., inputs selective for the attended feature will be enhanced while inputs selective for dissimilar features will be suppressed), its net effect on inputs' strength and low-frequency power within that region would be "null" or very small.

A puzzling finding in our experiments was the lack of attentional modulation in the θ band. Previous studies found that the phase of low-frequency oscillations modulates the strength of γ-band synchronization (Lakatos et al., 2005; Canolty et al., 2006), and it has been suggested that the cross-frequency phase synchronization between θ and γ oscillations play a fundamental role in memory processes (Sauseng et al., 2009, 2010), and learning (Tort et al., 2009). It is possible that in our scenario the power in the θ band was "preserved" by the attentional mechanism to facilitate the implementation of these task components. This hypothesis, however, remains to be tested.

References

Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, Panzeri S (2008) Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. J Neurosci 28:5696–5709.

Berens P, Keliris GA, Ecker AS, Logothetis NK, Tolias AS (2008) Feature selectivity of the gamma-band of the local field potential in primate primary visual cortex. Front Neurosci 2:199–207.

Bichot NP, Rossi AF, Desimone R (2005) Parallel and serial neural mechanisms for visual search in macaque area V4. Science 308:529–534.

Born RT, Bradley DC (2005) Structure and function of visual area MT. Annu Rev Neurosci 28:157–189.

Buschman TJ, Miller EK (2009) Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. Neuron 63:386–396.

Busse L, Katzner S, Treue S (2008) Temporal dynamics of neuronal modulation during exogenous and endogenous shifts of visual attention in macaque area MT. Proc Natl Acad Sci U S A 105:16380–16385.

Buzsáki G (2002) Theta oscillations in the hippocampus. Neuron 33:325–340.

Buzsáki G (2006) Rhythms of the brain. New York: Oxford UP.

Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT (2006) High gamma power is phase-locked to theta oscillations in human neocortex. Science 313:1626–1628.

De Baene W, Vogels R (2009) Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. Cereb Cortex. Advance online publication. Retrieved December 27, 2009. doi:10.1093/cercor/bhp277.

Felleman DJ, Kaas JH (1984) Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. J Neurophysiol 52:488–513.

Frien A, Eckhorn R, Bauer R, Woelbern T, Kehr H (1994) Stimulus-specific fast oscillations at zero phase between visual areas V1 and V2 of awake monkey. Neuroreport 5:2273–2277.

Frien A, Eckhorn R, Bauer R, Woelbern T, Gabriel A (2000) Fast oscillations display sharper orientation tuning than slower components of the same recordings in striate cortex of the awake monkey. Eur J Neurosci 12:1453–1465.

Fries P (2009) Neuronal gamma-band synchronization as a fundamental process in cortical computation. Annu Rev Neurosci 32:209–224.

Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. Science 291:1560–1563.

Fries P, Womelsdorf T, Oostenveld R, Desimone R (2008) The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. J Neurosci 28:4823–4835.

Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) High-frequency, long-range coupling between prefrontal and visual cortex during attention. Science 324:1207–1210.

Henrie JA, Shapley R (2005) LFP power spectra in V1 cortex: the graded effect of stimulus contrast. J Neurophysiol 94:479–490.

Katzner S, Nauhaus I, Benucci A, Bonin V, Ringach DL, Carandini M (2009) Local origin of field potentials in visual cortex. Neuron 61:35–41.

Kayser C, König P (2004) Stimulus locking and feature selectivity prevail in complementary frequency ranges of V1 local field potentials. Eur J Neurosci 19:485–489.

Kelly SP, Lalor EC, Reilly RB, Foxe JJ (2006) Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. J Neurophysiol 95:3844–3851.

Kelly SP, Gomez-Ramirez M, Foxe JJ (2009) The strength of anticipatory spatial biasing predicts target discrimination at attended locations: a high-density EEG study. Eur J Neurosci 30:2224–2234.

Khawaja FA, Tsui JM, Pack CC (2009) Pattern motion selectivity of spiking outputs and local field potentials in macaque visual cortex. J Neurosci 29:13702–13709.

Khayat PS, Spekreijse H, Roelfsema PR (2006) Attention lights up new object representations before the old ones fade away. J Neurosci 26:138–142.

Khayat PS, Niebergall R, Martinez-Trujillo JC (2010) Attention differentially modulates similar neuronal responses evoked by varying contrast and direction stimuli in area MT. J Neurosci 30:2188–2197.

Kreiman G, Hung CP, Kraskov A, Quiroga RQ, Poggio T, DiCarlo JJ (2006) Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. Neuron 49:433–445.

Kruse W, Eckhorn R (1996) Inhibition of sustained gamma oscillations (35–80 Hz) by fast transient responses in cat visual cortex. Proc Natl Acad Sci U S A 93:6112–6117.

Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol 94:1904–1911.

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Oscillatory entrainment as a mechanism of attentional selection. Science 320:110–113.

Leopold DA, Murayama Y, Logothetis NK (2003) Very slow activity fluctuations in monkey visual cortex: implications for functional brain imaging. Cereb Cortex 13:422–433.

Liu J, Newsome WT (2006) Local field potential in cortical area MT: stimulus tuning and behavioral correlations. J Neurosci 26:7779–7790.

Logothetis NK (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci 23:3963–3971.

Logothetis NK, Wandell BA (2004) Interpreting the BOLD signal. Annu Rev Physiol 66:735–769.

Martínez-Trujillo J, Treue S (2002) Attentional modulation strength in cortical area MT depends on stimulus contrast. Neuron 35:365–370.

Martinez-Trujillo JC, Treue S (2004) Feature-based attention increases the selectivity of population responses in primate visual cortex. Curr Biol 14:744–751.

Mitzdorf U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. Physiol Rev 65:37–100.

Mitzdorf U (1987) Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. Int J Neurosci 33:33–59.

Nauhaus I, Busse L, Carandini M, Ringach DL (2009) Stimulus contrast modulates functional connectivity in visual cortex. Nat Neurosci 12:70–76.

Reynolds JH, Heeger DJ (2009) The normalization model of attention. Neuron 61:168–185.

Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. Neuron 26:703–714.

Sauseng P, Klimesch W, Heise KF, Gruber WR, Holz E, Karim AA, Glennon M, Gerloff C, Birbaumer N, Hummel FC (2009) Brain oscillatory substrates of visual short-term memory capacity. Curr Biol 19:1846–1852.

Sauseng P, Griesmayr B, Freunberger R, Klimesch W (2010) Control mechanisms in working memory: a possible function of EEG theta oscillations. Neurosci Biobehav Rev 34:1015–1022.

Schroeder CE, Lakatos P (2009) The gamma oscillation: master or slave? Brain Topogr 22:24–26.

Sclar G, Maunsell JH, Lennie P (1990) Coding of image contrast in central visual pathways of the macaque monkey. Vision Res 30:1–10.

Seidemann E, Newsome WT (1999) Effect of spatial attention on the responses of area MT neurons. J Neurophysiol 81:1783–1794.

Siegel M, Donner TH, Oostenveld R, Fries P, Engel AK (2008) Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. Neuron 60:709–719.

Taylor K, Mandon S, Freiwald WA, Kreiter AK (2005) Coherent oscillatory activity in monkey area v4 predicts successful allocation of attention. Cereb Cortex 15:1424–1437.

Thut G, Nietzel A, Brandt SA, Pascual-Leone A (2006) Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. J Neurosci 26:9494–9502.

Tiesinga P, Sejnowski TJ (2009) Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? Neuron 63:727–732.

Tort AB, Komorowski RW, Manns JR, Kopell NJ, Eichenbaum H (2009) Theta-gamma coupling increases during the learning of item-context associations. Proc Natl Acad Sci U S A 106:20942–20947.

Treue S, Maunsell JH (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. Nature 382:539–541.

Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399:575–579.

Worden MS, Foxe JJ, Wang N, Simpson GV (2000) Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. J Neurosci 20:RC63(1–6).

Zeki S (1980) The response properties of cells in the middle temporal area (area MT) of owl monkey visual cortex. Proc R Soc Lond B Biol Sci 207:239–248.

2.5 Feature-based attention influences contextual interactions during motion repulsion

The perceived angle between objects moving in two different directions of motion is larger than the physical angle. This misperception is known as motion repulsion. In the present study we used this effect to test whether spatial interactions among sensory information can be modulated by feature-based attention.

Human subjects were asked to judge the direction of motion of two spatially non-overlapping random dot patterns (RDPs). One of the RDPs was composed of dots uniformly moving in a single direction and positioned at the center of gaze. The second RDP had the shape of an annulus. It was composed of two superimposed groups of dots moving in orthogonal directions and surrounded the central RDP.

When subjects judged only the motion direction of the central RDP, no motion repulsion was observed, since the influence of either of the two directions of motion in the surround RDP was counterbalanced. However, instructing subjects to attend to one particular motion direction in the surround RDP and simultaneously judge the motion direction of the central RDP, induced motion repulsion.

This result demonstrates that global feature-based attention can selectively change the relative influence of motion signals from non-overlapping spatial locations.

Available online at www.sciencedirect.com
ScienceDirect
Vision Research 46 (2006) 3651–3658
Vision Research
www.elsevier.com/locate/visres

Feature-based attention influences contextual interactions during motion repulsion

Tzvetomir Tzvetanov [a,*], Thilo Womelsdorf [b], Robert Niebergall [a], Stefan Treue [a]

[a] Cognitive Neuroscience Laboratory, German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany
[b] F.C. Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands

Received 20 January 2006; received in revised form 26 May 2006

Abstract

Visual perception is strongly shaped by the spatial context in which stimuli are presented. Using center-surround configurations with oriented stimuli, recent studies suggest that voluntary attention critically determines which stimuli in the surround affect the percept of the central stimulus. However, evidence for attentional influences on center-surround interactions is restricted to the spatial selection of few among several surround stimuli of different orientations. Here, we extend these insights of center-surround interactions to the motion domain and show that the influence of surround information is critically shaped by feature-based attention. We used motion repulsion as an experimental test tool. When a central target motion was surrounded by a ring of motion, subjects misperceived the direction of the foveal target for particular center-surround direction differences (repulsion condition). Adding an appropriate second motion in the surround counterbalanced the effect, eliminating the repulsion. Introducing feature-based attention to one of the two superimposed directions of motion in the surround reinstated the strong contextual effects. The task relevance of the attended surround motion component effectively induced a strong motion repulsion on the foveally presented stimulus. In addition, the task relevance of the foveal stimulus also induced motion repulsion on the attended surround direction of motion. Our results show that feature-based attention to the surround strongly modulates the veridical perception of a foveally presented motion. The observed attentional effects reflect a feature-based mechanism affecting human perception, by modulating spatial interactions among sensory information and enhancing the attended direction of motion.

Keywords: Motion repulsion; Feature-based attention; Contextual interactions

1. Introduction

Voluntary attention has a powerful influence on the control of contextual visual information (Gilbert, Ito, Kapadia, & Westheimer, 2000). Selective attention to stimuli surrounding a behavioral relevant stimulus can enhance the effective contrast of a central stimulus, or it can reduce perceptual sensitivity to the central stimulus (Zenger, Braun, & Koch, 2000; Freeman, Sagi, & Driver, 2001, 2003). Consistent with psychophysical evidence, physiological studies have revealed strong effects of attention on spatial interactions between center and surround stimuli in early visual cortical areas (Ito & Gilbert, 1999; Crist, Li, & Gilbert, 2001; Li, Piëch, & Gilbert, 2004). In these studies spatial attention modulated neuronal responses to oriented bars presented with different offsets and relative orientations in the center and surround of neuronal receptive fields.

While these studies reveal that spatial attention plays a pivotal role in structuring our visual environment by modifying the integration of nearby stimuli, they are limited in two respects. First, support for the role of attention is restricted to experiments using static bar or grating stimuli with different orientations. It is therefore unclear how spatial interactions in other visual domains such as motion is affected by voluntary attention.

* Corresponding author.
E-mail address: tzvetomir.tzvetanov@gmail.com (T. Tzvetanov).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.05.018

Second, modulation of center-surround interactions has exclusively been investigated with attentional selection based on spatial position. However, attention is also known to modulate visual perception based on the selection of feature information alone (Lankheet & Verstraten, 1995; Chen, Meng, Matthews, & Qian, 2005; Felisberti & Zanker, 2005). It remains unknown whether feature-based selection affects the spatial interactions of stimuli.

Here, we aim to shed light on these unresolved aspects by investigating feature-based attentional influences on center-surround interactions. Psychophysical studies have shown that attention can change the perceived direction of motion or enhance the perception of one motion among multiple presented ones. In these studies, attentional effects were observed for moving stimuli presented either in isolation (Chaudhuri, 1990) or as transparent surfaces containing superimposed direction of motion (Lankheet & Verstraten, 1995; Chen et al., 2005; Felisberti & Zanker, 2005). With transparent motion, attention to one direction of motion has been shown to reduce motion repulsion, i.e. the overestimation of the physical angular difference between two direction of motions is diminished (Marshak & Sekuler, 1979; Chen et al., 2005). While this finding reveals an influence of feature-based attention on motion repulsion, it does not show whether attention also modulates interactions of motion signals when they are spatially non-overlapping, similar to what has been observed in the orientation domain. Such effects of attention on spatial interactions has not been studied, even though motion stimuli presented in the surround are known to strongly influence motion processing in the center at the neuronal and behavioral level (Allman, Miezin, & McGuiness, 1985; Hiris & Blake, 1996; Kim & Wilson, 1997; Braddick, Wishart, & Curran, 2002). We therefore set out to test the influence of attention on spatial interactions in a motion repulsion paradigm with a center-surround stimulus configuration.

Subjects were asked to discriminate the direction of motion of a foveally presented target stimulus and/or the direction of motion of a parafoveally presented surround motion. Direction discrimination was characterized by parameters of psychometric functions. The experimental set-up contained five conditions: (1) a control condition with only a foveal stimulus, (2) a repulsion condition where a single motion in the surround along the leftward diagonal direction creates a misperception of upward motion of the foveal stimulus, (3) a no-repulsion condition with two superimposed motions of orthogonal directions in the surround, which was expected to result in no net effect of misperception of the central target stimulus, (4) an attentional control condition where subjects had to attend to one of the surround directions while the center motion was irrelevant, and (5) an attentional test condition where the subjects had to discriminate simultaneously motion in the center and the surround. These five conditions allowed us to disentangle the influence of feature-based attention in the surround on the perception of foveally presented motion.

2. Methods

A total of eight naive subjects participated in the study.[1] They had normal or corrected to normal vision and gave written consent for participating in the experiment.

2.1. Apparatus and stimuli

The experiment was conducted on a 21 inch CRT monitor at a refresh rate of 85 Hz and a resolution of 40 pixels per degree of visual angle, controlled by an Apple Macintosh G4 computer. Stimuli were random dot patterns (RDP) presented at the center of a white screen (luminance: 80.2 cd/m^2). Each dot extended 4×4 dark pixels (RDP absolute contrast of 22.6 cd/m^2). Dots moved within a circular or annular aperture at a speed of 8 degrees/sec in unidirectional or bidirectional translational motion. Upward motion was defined as zero degree, and leftward motions as negative values. The foveally presented target RDP had a radius of 1.5 degrees and contained 10 dots/deg^2. Its direction of motion was between ± 20 degrees of the vertical, sampled in one degree steps.

In four of the five conditions used, the target RDP was surrounded by an annular aperture (inner/outer radius: 1.5/6 degrees) as illustrated in Fig. 1B and C. For control condition C2, the surround annulus contained 100% coherent motion in either one of two possible directions with a 5 or 10 degrees offset relative to the -45 degrees diagonal (either $-55/-35$ or either $-50/-40$ degrees), with 10 dots/deg^2. For the remaining control and test conditions (Fig. 1C), two superimposed surfaces of moving dots were presented in the annular ring, with direction of motion along the leftward diagonal (-45 degrees ± 5 or either ± 10 degrees) in one surface and rightward diagonal motion in the second surface ($+45$ degrees ± 5 or either ± 10 degrees). Each surface contained 6 dots/deg^2.

2.2. Procedure

Subjects were seated in a dimly lit room 57 cm in front of the monitor. A chin rest was used to stabilize the head. They were instructed to fixate a small dark square centered on the screen. Trials were started by pressing the space bar on a computer keyboard, and 212 ms after the offset of the fixation square the stimulus was presented for 212 ms at the center of the screen. Two black lines, oriented at -45 degrees, were presented for 529 ms from the offset of the fixation square (positioned at about 7 degrees eccentricity, see Fig. 1B and C). They served as a reference for subjects judging the angular deviation of the motion in the annular surround task. The subjects had to report if the direction of motion of the small foveal RDP (the target) was to the left or right relative to his/her internal reference direction of upward motion by pressing corresponding keys on the computer keyboard. In conditions with surround task, they had to report if the motion in the leftward diagonal direction was moving more "counterclockwise" or "clockwise" from the diagonal formed by the reference lines by pressing corresponding keys.

The experiment included five conditions, four control conditions (referred to as C1–C4) and one test condition (Test) (see Fig. 1). Conditions were chosen to investigate the influence of the surround on motion discrimination of the central target, and the influence of attention on these interactions. In the first condition (C1) only the central target RDP was presented and the discrimination threshold and the perceived vertical direction was measured (cf. Fig. 1A). The second condition (C2) was designed to measure motion repulsion, i.e. the shift of the perceived direction of motion of the central target RDP. Subjects had to discriminate the

[1] From initially 9 subjects, one was excluded because that person did not reach criterion-level performance in the single surround task and the double-attentional task.

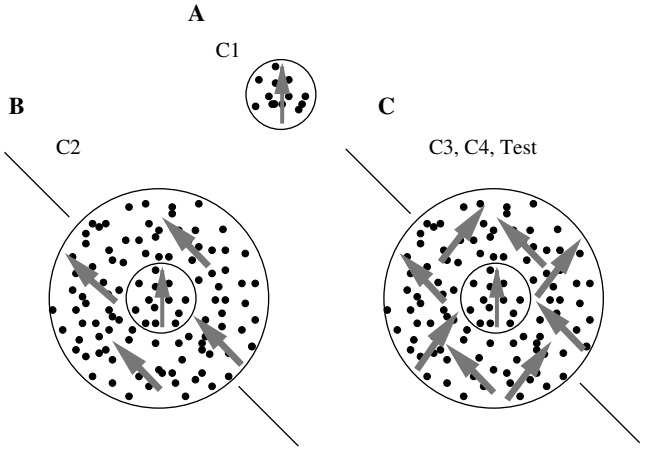


Fig. 1. Illustration of the stimuli in the different experimental conditions. The grey arrows show the global direction of motion in the corresponding part of the stimulus. (A) A foveal RDP containing nearly vertical upward motion; (B) the central target is surrounded by an annular aperture containing motion along the −45 degrees diagonal direction; (C) the target RDP is surrounded by an annular aperture containing two superimposed directions of motion along the two diagonals at ±45 degrees from the vertical upward direction.

direction of the target RDP in the presence of the annular surround moving at −45 degrees. The surround was behaviorally irrelevant since the subjects had to report only the motion direction of the central RDP target.

In the third control condition (C3) a second direction of motion was added in the annular surround, so that it contained superimposed motions with one motion surface moving in the leftward diagonal direction and the other surface moving in the rightward diagonal direction (+45 degrees). Subjects had to perform the direction discrimination task as in the previous conditions on the central target RDP with the motions in the surround being irrelevant. This condition was expected to result in no net effect of the surround motions on the perceived direction of an upward moving central RDP since the effects of the two surround motions are opposite and thus should cancel each other.

In control condition C4 the visual display and motion was identical to condition C3 but with a changed task. Subjects were instructed to judge only the leftward motion component in the annulus. They had to indicate whether that RDP moved "counterclockwise" or "clockwise" relative to the reference diagonal (cf. Fig. 1C). The direction of motions were adjusted either to −35/−55 or to −40/−50 degrees as a function of each subject's performance.

In order to direct attention to one of the surround direction of motions, we combined conditions C3 and C4 and required subjects to perform both task simultaneously. In this Test condition the visual display was identical to C3 and C4 (see Fig. 1C). Subjects had to judge the leftward direction in the surround, while at the same time they had to judge whether the central RDP moved clockwise or counterclockwise from upward. First they had to give the answer for the surround task, and then to the central target motion.

The control conditions were run in separate blocks of 100 trials, and the test condition was run in two blocks of 200 trials. All conditions were completed within two hours, and the experiment was repeated over two days, with the sequence of conditions randomized within a day and across subjects, with the exception that condition C4 was always measured before the Test condition and repeated until performance was between 60% and 90%.

To obtain the motion discrimination parameters for the central target RDP, a weighted staircase method was used for sampling the response curve of the subject (Kaernbach, 1991). Two staircases, with steps up/down of 3/1 degrees and 1/3 degrees (corresponding respectively to convergence points of 75% and 25%, see Kaernbach, 1991), were interleaved (Cornsweet, 1962). In addition, it avoided biases by having an equal number of right and left responses of the subjects. Feedback was provided to the subject for the overall performance on the surround task at the end of each corresponding experimental block. No feedback was used during the experimental blocks.

2.3. Data analysis

2.3.1. Parameter extraction of motion discrimination

A psychometric function was obtained for each experimental condition where the perceived target direction of motion of the central stimulus was measured. The psychometric function represents the proportions of "left" answers of the subject as a function of the target direction of motion. Using the maximum likelihood method together with the simplex algorithm for minimum search (Press, Teukolsky, Vetterling, & Flannery, 1997), each response curve was fitted with a logistic model of the form:

$$p(x) = \frac{1}{1 + \exp(-b(x - a))} \tag{1}$$

where x represents the direction of motion of the target RDP, $p(x)$ is the corresponding hit rate, a is the midpoint of the curve and b is related to its steepness. Thus, a is the direction for which subjects are equally likely to give a "left" or "right" response, i.e. the internal upward reference motion. Parameter b allows to compute the discrimination threshold defined as $\sigma = x_{p=84} - x_{p=50} = (1/b)\ln(21/4)$. It represents the magnitude of direction deviation in degrees that allows the subject to discriminate between the target direction relative to his/her internal vertical reference in 84% of the trials.

An example of staircase runs and a logistic fit to one response curve is shown in Fig. 2. Fig. 2A shows the two interleaved staircase runs, plotting the staircase trial number as a function of target direction of motion. Each staircase is starting at the opposite side from the convergence point, at an angular motion deviation from vertical of ±15 degrees. The staircases ensured a rapid convergence due to the asymmetric step sizes of the algorithms. Fig. 2B presents the pooled psychometric function, corresponding to the proportion of "left" responses as a function of target motion direction, together with the resulting maximum likelihood fit of the logistic

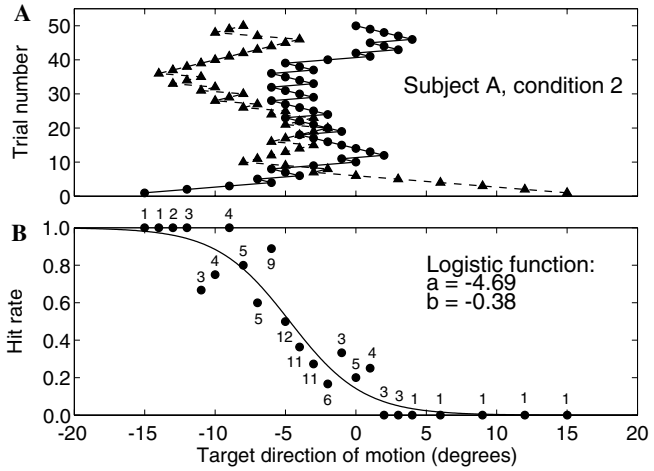


Fig. 2. Example of the resulting two staircase runs (triangles and circles) in (A), and the corresponding pooled experimental response curve (dots) together with the fitted logistic model (solid line) in (B). The example is from experimental condition C2 for subject A. In the bottom panel, the numbers above/below each data point correspond to the number of trials at this stimulus level. The bottom panel also illustrates the repulsion effect, i.e. the physical vertical direction of motion (0 degrees) is seen by the subject as moving to the right (about 15% "left" responses), and thus the midpoint is shifted to the left (negative value, $a = -4.69$).

function. The example illustrates the repulsion effect due to a surround direction of motion at −45 degrees, with the midpoint of the psychometric function shifted toward negative values (see Fig. 2).

All midpoint and threshold values of a given subject were obtained from psychometric functions containing 100 trials. The test condition had a total of 400 trials for each day. It was split in four consecutive 100 trials samples, and from each one a psychometric function fit was obtained, thus obtaining four test values per day (subsequently named T1–T4).

With regard to the discrimination of the surround motion (conditions C4 and Test), performance accuracy was computed as the percentage of correct responses for each 100 (400) trials run for C4 (Test) at each angular deviation of the surround. Then, the two points were used for extracting the parameters of the logistic function for the surround task by using the standard *logit* transform.

2.3.2. Statistical analysis

2.3.2.1. Center task. After conducting the experiment, one subject turned out to have particularly high thresholds for the center stimulus corresponding to very shallow psychometric functions in almost all test conditions. The corresponding staircase runs showed no consistent convergence properties. These individual results are presented in Appendix A but not used in the global analysis. Furthermore, two other subjects did not show the expected repulsion effect in condition C2 compared to conditions C1 and C3 (see Appendix A). This matches a report by Grunewald (2004) that about 1 out of 6 subjects does not show a motion repulsion effect. In addition, a previous study from our own laboratory showed that about 20–30% of the subjects provide small or no-repulsion effect reducing the strength of motion repulsion across the subject pool (cf. left panel in Fig. 4, Rauber & Treue, 1999). Nevertheless, these two subjects were included in the ANOVA (see below), and thus the data of 7 subjects were used.

We first conducted the Analysis of Variance on the full model with factors Experimental Condition (EC: 7 levels—C1–C3, T1–T4), Day of measurement (Day: 2 levels—D1, D2), and Subject as random factor. It did not show significant effects of Day of measurement and no statistical differences between the four test values (see Appendix B). For clarity, we therefore restrict the presentation of the results to the main effect of the experimental condition by pooling individual subject data across Day of measurement and conditions T1–T4. A restricted model was applied with only EC as main factor (one-way repeated measures ANOVA with 4 levels—C1–C3, Test).

2.3.2.2. Surround task. Since only two points were available for estimating the parameters of the psychometric function for the surround task, the *logit* transform ($\mathrm{logit}(x) = \ln((1-p)/p) = -b(x-a)$) could not be performed in those few cases where one of both of the data points was 0 or 1. The data of two subjects had to be discarded for this reason, and therefore the surround analysis included 6 data sets. As for the center task, Day of measurement did not show a significant effect (see Appendix B) and the results present the analysis once the data were pooled across days (paired *t*-test).

For the correlation analysis, data of 5 subjects for simultaneously center and surround were available, which provided 10 data points for the analysis (5 subjects × 2 center-surround pairs).

For the results from the individual subjects presented in the appendix, we obtained the 95% confidence intervals of each parameter using a parametric Bootstrapping method by simulating 2000 experimental runs (Efron & Tibshirani, 1993; Kaernbach, 1991). This involved the simulation of two interleaved staircases of 100 trials by assuming the experimental psychometric function as the theoretical curve and using the experimental parameters for the staircases (steps up/down of 1/3 and 3/1 degrees, starting values of ±15 degrees), then pooling the simulated data and fitting with a logistic model in the same way as in the experimental analysis. Data transformations and analyses were done with commercial software packages (Matlab, Mathworks Inc., MA; Prism, GraphPad Software, Inc., California).

3. Results

3.1. Motion repulsion

The location of the psychometric function reflects the subject's perceived upward direction, i.e. the motion direction for which the subject is equally likely to give a "left" or "right" response. A negative value shows that a subject's vertical reference is tilted to the "left", i.e. he/she reports a counterclockwise motion as vertical. In condition C2, it represents a repulsion since the leftward motion of the surround biases a leftward center motion to be perceived as vertical, i.e. further away from the surround direction. In other words, the physical vertical direction (0 degrees) appears to be moving to the right (cf. Fig. 2). The average results are shown in Fig. 3. Fig. 3A shows the location parameter as a function of the experimental condition. As expected, there was no significant shift of the perceived vertical in the absence of the surround stimulus (condition C1). With the introduction of the surround annulus with one direction of motion angled −45 degrees to the left, the perceived motion of the center stimulus (condition C2) was significantly shifted toward negative values, reflecting motion repulsion. Adding a second surface of motion in the surround (condition C3) abolishes the motion repulsion effect. Most importantly though, in the dual task, with attention to one of the two directions of motion in the surround (Test), motion repulsion reemerged. Thus, attending to one direction of the transparent motion in the surround resulted in motion repulsion in the center.

This pattern of results was statistically confirmed using an ANOVA (see Section 2.3). The ANOVA showed a highly significant effect of the experimental condition on the perceived vertical reference ($p < 0.01$). To obtain a complete statistical overview, we performed multiple comparison tests with Bonferroni corrected significance levels. We found significant differences between: C1 and C2, C1 and Test, C3 and C2, C3 and Test ($p < 0.05$). Thus, the two conditions where we did not expect a bias of motion perception (C1 and C3) were both statistically different from the two conditions where we expected a bias (C2 and Test).

Fig. 3B shows the means of the location parameter in the two surround conditions (C4 and Test). The C4 value near −45 degrees shows that the subjects correctly perceive the −45 degrees diagonal reference direction when performing the single surround task, i.e. when the center stimulus and surround rightward diagonal motion were task irrelevant. Asking the subjects to simultaneously perform the center task (Test condition) showed a strong effect on the perceived surround diagonal direction, with a repulsion effect of about 8 degrees (note that the repulsion effect in the surround is opposite to that in the center, with location parameter being positively shifted). This pattern of results was statistically confirmed with a restricted model to Experimental Condition (paired *t*-test, $p < 0.01$).

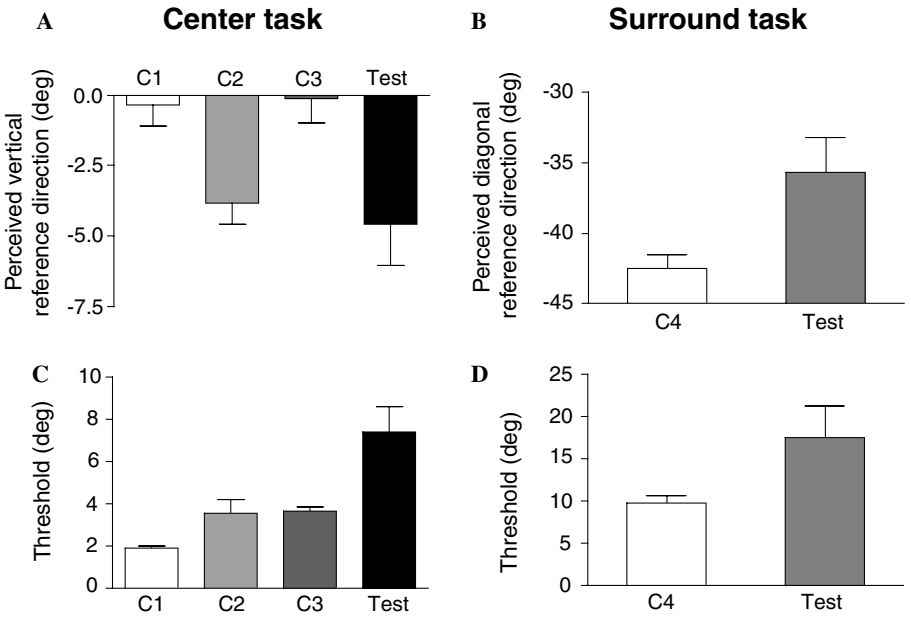


Fig. 3. (A and B) Mean values of perceived vertical (A) and diagonal (B) reference directions as a function of the experimental condition. (C and D) Mean values for vertical (C) and diagonal (D) discrimination thresholds as a function of the experimental condition. For the center task, the reference direction of motion was the vertical upward motion (defined as 0 degrees). For the surround task, the reference direction of motion was the diagonal at −45 degrees. Error bars are standard errors (center task, $n = 7$; surround task, $n = 6$). The data presentation is restricted to the statistically significant effect of experimental condition (see Section 2.3).

3.2. Discrimination performance

The psychometric functions did not only provide estimates of the perceived vertical or diagonal reference directions, but also of the discrimination thresholds (σ). These thresholds describe the deviation of motion direction for which subjects reported the correct response in 84% of the trials. They reflect the difficulty of discriminating two close direction of motions, with higher values showing worse discrimination ability of the subjects. Fig. 3C and D present the average thresholds for each experimental condition. For the center task (Fig. 3C), the restricted model ANOVA showed a significant effect of experimental condition ($p < 0.01$). Multiple comparisons (Bonferroni corrected) showed that thresholds are significantly higher in the Test condition compared to the control conditions (C1–C3) ($p < 0.01$). For the surround task (Fig. 3D), the experimental condition similarly showed a significant effect (paired t-test, $p < 0.05$), with thresholds being higher in the Test condition.

3.3. Center-surround performance trade-off

We assessed any possible relation between center and surround performance by computing the correlation coefficients of each pair from the four measured variables in the test condition: center location, surround location, center threshold, surround threshold. One pair was significantly correlated, center location × center threshold ($r = -0.84$, $p < 0.01$). The remaining five were not significant: surround location × surround threshold ($r = 0.45$, $p = 0.20$), center location × surround location ($r = -0.21$, $p = 0.55$), center location × surround threshold ($r = 0.29$, $p = 0.41$), surround location × center threshold ($r = 0.02$, $p = 0.95$), and center threshold × surround threshold ($r = -0.45$, $p = 0.20$). Thus, only the parameters of the central task were correlated among themselves, with no trade off between performance of center and surround task.

4. Discussion

This study investigated feature-based attentional effects on center-surround interactions during motion processing. Our results show that feature-based attention strongly biases the processing of surround and center information.

In experiment C2, we first observed that a single moving stimulus in the surround, at a direction of −45 degrees of a central motion signal, induced the classical motion repulsion effect: subjects misperceived the direction of motion of a foveally presented RDP, i.e. they overestimated the angle between the directions of motions in the center and its spatial surround. In condition C3, we showed that motion repulsion is abolished by presenting a second motion signal in the surround moving at +45 degrees offset from the foveal one, thus appropriately counterbalancing the motion direction at −45 degrees in the surround. In the Test condition, we instructed subjects to attend to only one of the two surround signals. Importantly, this reinstated the repulsion effect observed in C2. This shows that feature-based attention increases the influence of the attended motion direction, and/or reduces the influence of the unattended motion direction in the surround. Noteworthy, the

effect of attention in our task is behaviorally detrimental, because it induces a distortion of the perceived direction of motion of the foveally presented stimulus.

Our finding of an attentional enhancement of the repulsion effect complements and extends a recent finding by Chen et al. (2005). These authors used a single foveally presented stimulus with two superimposed directions of motion moving at an angular deviation of 45 degrees. When subjects attended one direction of motion in order to detect a speed change, the motion repulsion for this direction was reduced. This finding suggests that feature-based attention selectively changes the relative influence of motion signals moving across each other. Consistent with this findings we report that feature-based selection of one of two superimposed motion directions in the spatial surround enhances its influence on the center. While the attentional effect in this previous study is a reduction of motion repulsion, in our study feature-based attention selected the motion signal that induced the net repulsion effect, and hence motion repulsion was reinstated. Taken together, the results of both studies complement each other by showing that feature-based selection within transparent surfaces modulates the processing of motion components, thereby changing motion repulsion and our perception of visual motion.

The similar conclusions in Chen et al. (2005) study and in our experiment are noteworthy not only because they were obtained with different experimental paradigms involving different stimulus layouts. In particular, while we derived the misperception of the perceived vertical reference direction from psychometric functions of discrimination performance, Chen et al. (2005) derived it by requiring subjects to manually adjust the direction of motion relative to a previously shown reference direction. Despite differences in experimental designs, both studies come to the same conclusion about the influence of attention to modulate contextual interactions.

Interestingly, we observed motion repulsion not only for the foveal target stimulus but also for the perceived diagonal direction of the surround motion in the dual task condition with attention to the center and to the surround. Thus, directing attention to motion in the center and to motion in the spatial surround mutually enhances the relative influence of attended motion directions for both tasks: Attended surround motion induces motion repulsion of the center motion, while attended center motion biases the perception of the surround motion. This finding was unexpected because the size of the center stimulus was small compared to the annular surround stimulus (but cf. Kim & Wilson, 1997). Importantly though, the amount of motion repulsion observed for center and surround discrimination was not related to each other and thus does not affect our finding of an influence of feature-based attention in the spatial surround on foveal motion perception.

In contrast to previous studies on the effect of attention on motion processing, we focused on spatial interactions among motion signals. In this respect, our finding extends reports from the orientation domain, which demonstrated that attention biases perception by modulating spatial interactions between stimuli of different orientations (Ito & Gilbert, 1999; Freeman et al., 2001; Crist et al., 2001; Li et al., 2004). Consistent with these findings we show that attention can cause biases in perceived motion directions.

The attentional modulation observed in our study relies on feature-based selection of motion signals in the surround, because surround directions of motion spatially overlapped. This finding shows that feature-based attention critically determines which motion directions in the surround influence the processing of the foveal target stimulus. Consistently, neurophysiological studies have shown that feature-based attention modulates neuronal responses in motion sensitive cortical areas (Treue & Martinez-Trujillo, 1999; Martinez-Trujillo & Treue, 2004). These studies suggest that feature-based attention most strongly enhances the response gain of neurons with a tuning preference for the attended motion direction, and decreases the response of neurons preferring directions of motion offset from the attended direction (Martinez-Trujillo & Treue, 2004). Applied to our experiment, this effect explains the enhanced influence of the attended contextual motion signal based on attentional gain modulation.

In addition to the selective feature-based enhancement of one of two motion directions in the surround, attention could have exerted a direct influence on the spatial integration process itself. However, our task does not allow to disentangle a possible attentional modulation of the spatial integration from the observed attentional enhancement of surround motion signals. It awaits to be seen in future studies whether attention modifies not only the gain of surround motion signals but also the nature of the spatial interactions among center and surround processes.

In summary, our study demonstrates that attention modulates the contextual interactions between motion signals and affects the perceived direction of motion in a motion repulsion paradigm. The consistency of our findings with previous psychophysical and neurophysiological reports studying motion and orientation suggests that attentional modulation of contextual interactions could be a general principle deployed by the visual system.

Acknowledgment

This study was supported by funding from the Bernstein Center for Computational Neuroscience, Göttingen (Grant No.: 01GQ0433).

Appendix A. Individual results for the center task

Fig. A.1 shows the perceived vertical direction (location parameter) for each subject and each experimental condition, from the two measurements at successive days (left/right panels for each subject). For each subject, the control condition C1 is plotted as a grey band. Its height represents the 95% confidence interval (CI). Five of the subjects

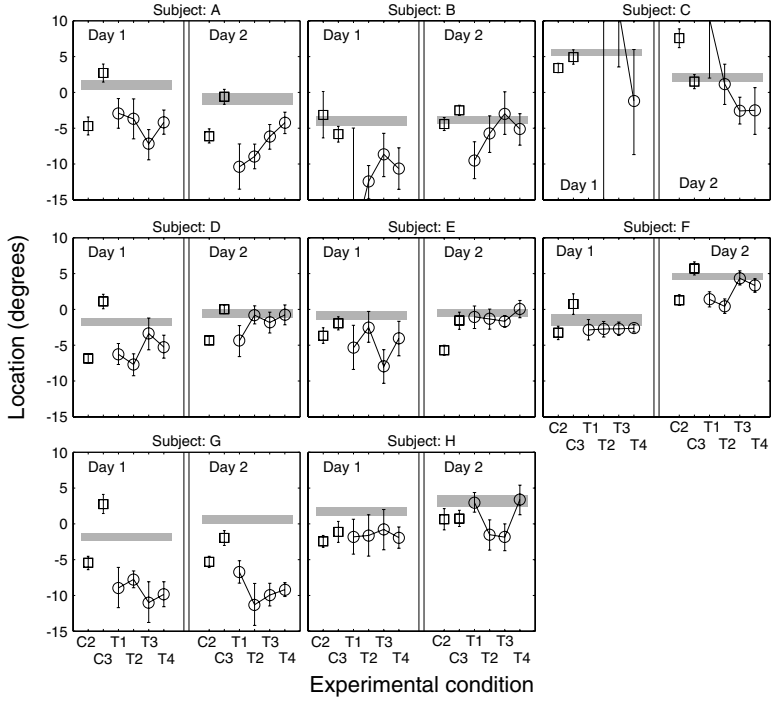


Fig. A.1. Perceived vertical reference direction of motion (Location parameter, *a*) as a function of the experimental condition, for each day and subject. Error bars are 95% confidence interval. Grey rectangles represent 95% confidence intervals for the C1 condition. Negative values on the ordinate indicate that perceived vertical reference is to the "left" from the true vertical direction of motion i.e., reflecting motion repulsion. The points out of range of the plots correspond to (location [95%CI]): subject B, day 1—T1 = −21.1 [−45.3;−5.0]; subject C, day 1—T1 = 14.7 [10.9;18.6], T2 = 24.8 [−28.6;49.1], T3 = 11.5 [3.5;20.2], day 2—T1 = 10.2 [2.0;18.4].

(subjects A, D–G) showed clearly a motion repulsion effect in conditions C2, i.e. the locations are shifted toward negative values compared to conditions C1 and C3. There were two subjects (B and H) who did not show a repulsion effect in condition C2. However, similar inter-subject variability for motion repulsion has been reported before (Grunewald,

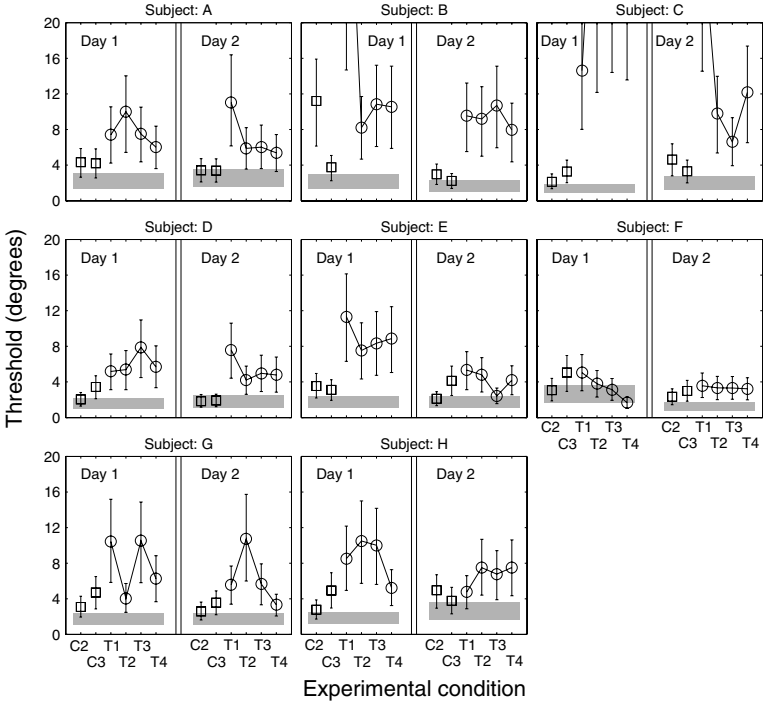


Fig. A.2. Discrimination thresholds for each subject as a function of the experimental condition and day of measurement. Same format as Fig. A.1, but with discrimination threshold on the *Y*-axis. The points out of range of the plots correspond to (location [95%CI]): subject B—T1 = 45.7 [14.7;1039.5] ; subject C, day 1—T1 = 14.6 [8.0;23.3], T2 = 43.9 [12.2;+∞], T3 = 31.2 [14.4;6.9], T4 = 28.2 [13.6;48.2], day 2—T1 = 30.4 [14.5;57.5].

2004). These individual results also show that attention to one motion direction in the surround (T1–T4) re-evoked motion repulsion compared to condition C3, in which the same visual motion in the spatial surround was present but no-repulsion effect was observed.

To inspect the performance differences across subjects for the center task in more detail, we present the threshold levels for each subject and experimental run in Fig. A.2. In addition, this plot documents the poor performance of subject C, with high discrimination thresholds in all test conditions. Interestingly, the two subjects (B and H) with no effect of motion repulsion in condition C2 showed the same pattern of low threshold for condition C2 compared to C1 and C3, exactly as the remaining subjects. This excludes the possibility that it is a poor discrimination performance of these two subjects which led to the lack of motion repulsion.

Appendix B. Analysis of variance

We performed a full model analysis of variance on the center and surround task parameters, with fixed factors experimental condition (EC), day of measurement (Day), and "subject" as random variable.

B.1. Center task

For the location parameter, the perceived vertical reference direction of motion, the ANOVA showed a highly significant effect of EC on the perceived vertical reference ($p < 0.01$), no effect of Day ($p = 0.27$), and no interaction between the two factors ($p = 0.44$). Furthermore, the ANOVA reported a significant "subject" effect ($p < 0.01$) which we attribute to the mentioned inter-subject variability in the amount of motion repulsion. For the discrimination thresholds, the ANOVA showed a significant effect of experimental condition ($p < 0.01$), no effect of day of measurement ($p = 0.16$) and no interaction effect between the two factors ($p = 0.38$).

B.2. Surround task

For the location parameter representing the perceived diagonal reference direction, the ANOVA showed a significant effect of EC on the perceived leftward-diagonal reference ($p < 0.05$), no effect of Day ($p = 0.43$), and no interaction between the two factors ($p = 0.88$). For the discrimination thresholds, the 2-way ANOVA showed a significant effect of EC ($p < 0.05$), no effect of Day ($p = 0.46$), and no interaction pattern between the two factors ($p = 0.74$).

References

Allman, J., Miezin, F., & McGuiness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception, 14*, 105–126.

Braddick, O., Wishart, K., & Curran, W. (2002). Directional performance in motion transparency. *Vision Research, 42*, 1237–1248.

Chaudhuri, A. (1990). Modulation of the motion aftereffect by selective attention. *Nature, 344*, 60–62.

Chen, Y., Meng, X., Matthews, N., & Qian, N. (2005). Effects of attention on motion repulsion. *Vision Research, 45*, 1329–1339.

Cornsweet, T. (1962). The staircase method in psychophysics. *American Journal of Psychology, 75*, 485–491.

Crist, R., Li, W., & Gilbert, C. (2001). Learning to see: experience and attention in primary visual cortex. *Nature Neuroscience, 4*, 519–525.

Efron, B., & Tibshirani, R. (1993). *An introduction to the Bootstrap. Monographs on statistics and applied probability*. 0-412-04231-2. Chapman & Hall/CRC.

Felisberti, F. M., & Zanker, J. (2005). Attention modulates perception of transparent motion. *Vision Research, 45*, 2587–2599.

Freeman, E., Driver, J., Sagi, D., & Zhaoping, L. (2003). Top-down modulation of lateral interactions in early vision: does attention affect integration of the whole or just perception of the parts? *Current Biology, 13*, 985–989.

Freeman, E., Sagi, D., & Driver, J. (2001). Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers. *Nature Neuroscience, 4*, 1032–1036.

Gilbert, C., Ito, M., Kapadia, M., & Westheimer, G. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Research, 40*, 1217–1226.

Grunewald, A. (2004). Motion repulsion is monocular. *Vision Research, 44*, 959–962.

Hiris, E., & Blake, R. (1996). Direction repulsion in motion transparency. *Visual Neuroscience, 13*, 187–197.

Ito, M., & Gilbert, C. (1999). Attention modulates contextual influences in the primate visual cortex of alert monkeys. *Neuron, 22*, 593–604.

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics, 49*, 227–229.

Kim, J., & Wilson, H. (1997). Motion integration over space: interaction of the center and surround motion. *Vision Research, 37*, 991–1005.

Lankheet, M. J., & Verstraten, F. A. J. (1995). Attentional modulation of adaptation to two-component transparent motion. *Vision Research, 35*, 1401–1412.

Li, W., Piëch, V., & Gilbert, C. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience, 7*, 651–657.

Marshak, W., & Sekuler, R. (1979). Mutual repulsion between moving visual targets. *Science, 205*, 1399–1401.

Martinez-Trujillo, J., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology, 14*, 744–751.

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (1997). *Numerical Recipes in C: The art of scientific computing*. Cambridge University Press.

Rauber, H.-J., & Treue, S. (1999). Revisiting motion repulsion: evidence for a general phenomenon ? *Vision Research, 39*, 3187–3196.

Treue, S., & Martinez-Trujillo, J. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399*, 575–579.

Zenger, B., Braun, J., & Koch, C. (2000). Attentional effects on contrast detection in the presence of surround masks. *Vision Research, 40*, 3717–3724.

2.6 Contribution of spike timing to contrast and motion direction coding by single neurons in macaque area MT

The average firing rate of a neuron over a predefined time interval is typically used to correlate the cell's response to behavior. However, the majority of single cell studies intending to do that in awake, behaving monkeys have found that neurons systematically underperform the animals' behavioral abilities. Two possible reasons for this result are, first, that the brain uses population codes, pooling the activity of many neurons in order to increase the signal-to-noise ratio and therefore improve coding and performance. Second, that single neurons use additional sources of information such as the timing of spikes within a spike train in addition to the firing rate alone, which also improves coding.

In the present study we tested whether spike time can be used as an additional source of information to improve the performance of single neurons. Different from previous studies addressing this issue in MT neurons, we used the metric analysis proposed by Victor and Purpura (1996) and tested whether train metrics that incorporate the temporal structure of spiking activity lead to higher information content than spike counts. We further compared the performance of an ideal observer model, using spike-time-based metrics and spike counts (produced by MT neurons) at discriminating the contrast and the direction of moving stimuli. We found that for stimuli that varied in contrast and direction, the metric D^{spike}, which pools information about spike counts and spike timing, led to higher information content, and performance by the ideal observer model than spike counts alone.

Our results demonstrate that spike timing is an additional source of information that can be used by single neurons in area MT to improve performance during direction and contrast discrimination tasks relative to using spike counts alone.

Contribution of spike timing to contrast and motion direction coding by single neurons in macaque area MT

Running head: Spike time coding in area MT

Authors: Adam J. Sachs[1,2], Paul S. Khayat[1], Robert Niebergall[1,3] and Julio C. Martinez-Trujillo[1]

Author addresses: All three authors have the same address as corresponding author

Affiliations: 1. Cognitive Neurophysiology Laboratory, Department of Physiology, McGill University

2. Division of Neurosurgery, The University of Ottawa

3. Cognitive Neuroscience Laboratory, German Primate Center

Corresponding author: Julio C. Martinez-Trujillo

McIntyre Medical Sciences Building

Department of Physiology

3655 Promenade Sir William Osler, Room 1223

Montreal, PQ H3G 1Y6

julio.martinez@mcgill.ca

(514)398-6024

Number of figures: 11

Number of tables: 0

Contents of supplemental material: Supplementary materials section (5 pages), 2 Supplementary materials figures

Number of pages: 49

Abstract

We examined spike trains produced by neurons in area MT of two awake monkeys during the presentation of moving stimuli to determine if the timing of the spikes carries information about the stimulus direction and contrast. We assessed the performance of three spike train metrics, D^{spike} and $D^{product}$ (in which spike timing is relevant to coding), and a rate code metric, D^{count} (in which spike timing is irrelevant), on the recorded data, and on computer generated spike trains. In two different experiments, moving random dot patterns with different directions and contrasts were presented inside the cells' receptive field. We analyzed data during the first 200 milliseconds of stimulus presentation from 205 neurons in two animals. In each neuron, we computed the information entropy of clustering to the stimuli with varying contrast and direction using the three metrics. We concentrated on pairs of stimuli for which a theoretical observer analysis using D^{count} yielded a discrimination performance of $0.55<P_{correct}<0.82$. For both stimulus attributes, the large majority of neurons showed the highest clustering entropy using D^{spike}, followed by $D^{product}$, and D^{count}. This was corroborated by the best performance of a theoretical observer model at discriminating different contrasts and directions using D^{spike}. Our results suggest that the spike timing of MT neurons, quantified by the metric D^{spike}, is a source of available information that could be used by the visual system to improve performance during direction and contrast discrimination tasks.

Keywords: Neuron, area MT, motion, contrast, spike trains, temporal coding

Introduction

Single neurons in the primate visual cortex can encode different stimulus features through differences in their firing rates. For example, in area MT of macaques, tuning curves can be generated from variations in a neuron's mean firing rate in response to changes in the direction of a moving stimulus presented inside the cell's receptive field (RF) (Dubner and Zeki 1971). Although such tuning curves seem to consistently follow the same profile, there is significant variability around the mean firing rate across individual presentations of the same motion direction (Buracas et al. 1998; Softky and Koch 1993). This variability may be associated with the highly irregular organization and seemingly unpredictable patterns of clustering exhibited by the neurons' spike trains (Holt et al. 1996; Hu et al. 2002; Rodieck et al. 1962; Shadlen and Newsome 1998).

One interpretation of irregular spike timing is that it is an inevitable consequence of random biological factors, and represents internal noise. Individual units acting as independent detectors or discriminators of a signal embedded in noise could, theoretically, boost their reliability by pooling their responses (Green and Luce 1975; Watson 1990). However, adjacent neurons in cortical areas such as MT are most likely not independent. Local field potentials (LFPs) seem to aggregate neuronal activity with a spatial extent on the order of 500 μm (Kruse and Eckhorn 1996; Liu and Newsome 2006), and anatomical studies have revealed complex networks of intrinsic and extrinsic interconnections among neurons (Malach et al. 1997).

A different interpretation, however, is that spike trains produced by sensory neurons in response to stimulus variations contain some degree of temporal structure that is informative about the stimulus. For example, Reich and coworkers demonstrated that the temporal structure of spike trains produced by V1 neurons carry information about the contrast of a stimulus (Reich et al. 2001b). Thus, it is likely that the same phenomenon occurs in other visual areas downstream from V1, such as area MT. This may be plausible for several reasons. Firstly, simple cells in V1, an important input to MT (Born and Bradley 2005), have receptive fields with interposed on and off regions (Hubel and Wiesel 1968) and periodic output to a motion stimulus (Carandini et al. 1997; Hubel and Wiesel 1968; Movshon et al. 1978). Secondly, power spectral analysis of local field potentials in area MT shows tuning in the gamma band region (Liu and Newsome 2006), which could be reflected in the spike trains. Thirdly, intracellular voltage oscillations have been shown to influence spike

timing in intracellular recordings from area MT neurons in macaques (Fellous et al. 2001; Fellous et al. 2004; Hunter et al. 1998).

Previous studies in area MT have examined whether the timing of the spikes contains information about the stimulus (Bair and Koch 1996; Buracas et al. 1998; Fellous et al. 2004; Osborne et al. 2004; Masse and Cook 2008). Using vector space-based analyses and related methods, they have reported that the information provided by a spike train about a stimulus direction is larger when considering spike times than when considering only the spike count (firing rate). Here, we expand this work by using a metric-based analysis and targeting the question of whether, under circumstances in which the firing rate is less informative (i.e. for stimuli lying relatively close along the slope of an MT neuron's tuning curve for contrast or direction), a theoretical observer model will benefit from using different sources of spike time information (e.g., periodic vs. aperiodic). Additionally, we examined whether individual features of the spike trains such as the latency of activity (relative to stimulus presentation) and distribution of interspike intervals are responsible for a potential coding advantage offered by spike time-based metrics.

In two separate experiments, we recorded responses of MT neurons in two macaques to stimuli with different motion directions and contrasts. We then determined the information entropy of clustering to the stimuli using different spike train metrics: D^{count}, D^{spike} (Victor and Purpura 1997), and a vector product metric, $D^{product}$ (Schreiber 2003). We additionally computed the performance of a theoretical observer at discriminating motion direction and contrast using each metric. We found that the temporal structure of spike trains fired by MT neurons contains additional information about stimulus direction and contrast relative to the firing rate alone. Such information seems to reflect the precise timing of individual spikes within the train rather than only the latency to the activity onset or the distribution of interspike intervals. Furthermore, we found that a theoretical observer improves its contrast and motion direction discrimination using spike timing information relative to when using only the firing rate.

Methods

1. Animal Preparation.

Two *Macaca mulatta* were used in the experiments. Each animal was surgically prepared for single unit recordings as previously described (Martinez-Trujillo and Treue 2004). Briefly, during each experimental session, the head was fixed with a surgically implanted head post, and eye movements were tracked with a video based eye tracker (Eyelink II, SR Research Ltd., Canada). During each experimental session, one or two penetrations were made using tungsten microelectrodes (FHC Inc., USA, impedance=0.5-3MΩ) and an electric microdrive (Plexon Inc., USA) positioned on the top of a recording chamber over the right parietal bone. Structural magnetic resonance images were used to localize the area of recording, and neurons were classified as MT units based on their response properties (receptive field size and direction tuning for linear motion direction and spiral stimuli) (Martinez-Trujillo and Treue 2004). Single unit responses were isolated and sorted online and offline using Plexon spike sorting software (Plexon Inc., USA), in order to extract the times of occurrence of each action potential. These protocols and procedures were in accordance with Canadian Council on Animal Care guidelines and approved by the McGill University Animal Care Committee.

2. Stimuli and Task.

The data reported in experiment 1 are part of a larger data set from another study (Khayat et al. 2008). Stimuli consisted of moving random dot patterns (RDPs) presented on a rear projection screen using a video projector (NEC WT610, NEC Inc., Japan). The screen resolution was 1024×768 with a refresh rate of 85 Hz. Each pixel was 0.083 cm^2 and each dot size was 0.17 cm^2. The monkey's eyes were 57 cm away from the monitor. At this distance, each pixel approximated 0.08 degrees of visual angle.

The dots within a pattern moved with 100% coherence at the preferred speed of the recorded neuron and had infinite lifetime. In different trials, the initial dots' position varied, although the RDP's dot density remained constant. This avoided using the same initial dots' position (seed) within a stimulus type, and different seeds across stimuli, which may produce specific patterns of spike timing corresponding to each stimulus type. When a dot crossed the aperture border, it was re-plotted on the

opposite side avoiding dots' appearance and disappearance at random positions within the aperture.

It is important to note that our stimuli are different from RDPs used in previous studies in area MT containing dots with less than 100% coherence and with random fluctuations in the dot statistics (e.g., Masse and Cook 2008). Such fluctuations can produce changes in human behavioral thresholds (Barlow and Tripathy 1997) and likely in the firing pattern of neurons in area MT. The stimuli used in both experiments of this study did not contain such fluctuations. Rather, they produced a stable percept of a rigid surface of approximately equaled spaced dots moving behind a circular aperture (Martinez-Trujillo and Treue 2002; Treue and Martinez-Trujillo 1999).

The animal had to fixate on a central fixation dot (0.3° square) centered in a circular fixation window (invisible to the animal) with diameter 1.5°, and initiate a trial by pressing a button. After 470 milliseconds (msec), two pairs of RDPs appeared - one pair located inside the RF of the recorded neuron, and the other located outside in the opposite hemifield (figure 1a). After a random delay (1480 - 3670 msec), the fixation dot changed luminance and the monkey had to release the button within a 350 msec window, starting 150 msec after the change, in order to obtain a reward (drop of water or juice). Trials in which the monkey broke fixation before releasing the button, or in which it released the button before the fixation dot changed luminance were aborted and considered errors. The animals correctly performed the task in more than 90% of the trials.

Each pair of RDPs consisted of one high contrast stimulus [13 cd/m^2 – using the standard deviation contrast method for RDPs (Moulden et al. 1990)] moving in the neuron's antipreferred (null) direction, and one test RDP (figure 1A). The latter could have two different configurations across trials: a) it could move in the neuron's preferred direction with trial to trial changes in contrast (100%, 14%, 1.5%, 0.7%, 0.3%, 0.1%, and 0.02% relative to the high contrast stimulus), or b) it could have the same contrast as the null pattern but with trial to trial variations in motion direction (spaced in intervals of 15° departing from the preferred direction until 90 degrees away from that direction). The direction and contrast values were chosen in such a manner that they produced similar response variations along the neurons' direction and contrast response functions [see figure 2, Martinez-Trujillo and Treue (2002), and Treue and Martinez-Trujillo (1999)].

In order to control for any potential impact that the stimulus configuration in experiment 1 might have had on temporal coding, we analyzed a second data set from another experiment (Niebergall et al. 2008). The task and timing of the trials was similar to experiment 1, however, here there was only one RDP inside the neurons RF. The direction of motion of the RDP changed between trials in steps of 30°, departing from the preferred direction (figure 1B). The two animals participating in this experiment were the same as in experiment 1. However, experiment 2 was conducted approximately one year later.

3. Spike train metrics

Our aim was to determine if area MT neurons could transmit information about stimulus attributes (in this case, contrast and direction) *via* a coding scheme based on spike timing beyond that of a simple rate code. To do this, we analyzed spike trains using the metric space method of Victor and Purpura (1997). Three metrics are considered: D^{count}, $D^{spike}(q)$, and one we refer to as $D^{product}(\sigma)$. These metrics produce measurements of the distance between spike trains. Distance can be measured in terms of the "cost" of transforming one spike train into the other. In D^{count} , the simplest of the three metrics, there is a cost of 1 for every spike added or removed, but no cost for moving a spike in time. This reduces to the arithmetic difference in firing rates (r_2-r_1), and thus implies a simple rate code in which the timing of spikes is irrelevant.

In $D^{spike}(q)$, there is a cost of 1 for adding or removing a spike (as in D^{count}), but there is also a cost, q, associated with moving a spike in time. The total "distance" from one spike train to the other is the sum of the cost associated with one of two elementary steps required to transform one spike train into the other: $q \cdot \Delta t$, the cost of shifting a spike in time, or 1, the cost of adding or deleting each spike. The parameter q is in sec^{-1}. The use of these two elementary steps is governed by the following rule: If two spikes in different spike trains are separated in time by an interval greater than 2/q (in seconds), then it is less costly to simply delete and replace the spike at the second location, and this operation is done at a cost of 2. Thus q reflects temporal precision since it can be thought of as the difference in the timing of occurrence of two spikes that makes just as much difference to the nervous system as a spike deletion (Di Lorenzo and Victor 2003). $D^{spike}(q \rightarrow 0) = D^{count}$, and therefore the first

point on the $D^{spike}(q)$ entropy function reduces to D^{count}. We used the spike toolkit (Goldberg et al. 2005) to calculate $D^{spike}(q)$ and the information entropy.

The third spike train metric is based on the vector product between spike trains (Schreiber 2003), and in keeping with the Victor and Purpura (1997) notation, we refer to it as $D^{product}$. Each spike train is convolved with a Gaussian filter, with size (kernel) σ, to form a spike density function s_i. We define the distance between two spike trains as:

$$D^{product}(\sigma) = 1 - \frac{\langle s_i, s_j \rangle}{\|s_i\| \cdot \|s_j\|},$$

where σ is the Gaussian filter size parameter. The intuition behind this metric is that spike trains are considered similar if the normalized product of their spike density functions is high.

4. Theoretical Observer model

In this section, we illustrate the strategy of a theoretical observer model subject to a metric based strategy. For simplicity, the analysis is restricted to a 2-alternative forced choice (2AFC) discrimination task. Consider two equiprobable signals, $S(\theta_0)$ and $S(\theta_1)$, where θ is the modulated stimulus parameter. One of these signals, $f(\theta_0,\theta_1)$, is presented to the animal and represented in the early visual system.

The information flow continues to area MT, where a neuron receiving the generated inputs (e.g., from area V1) produces a spike train, $r(f(\theta_0,\theta_1))$. We are asking the theoretical question "If a neuron or neural network located downstream only has access to this spike train, how accurately can it assign the signal to one of two categories corresponding to the different stimuli?" To address this question, we consider the performance of a theoretical system (either a neuron or network of neurons downstream) in which the spike train is compared to two spike train templates, $\bar{r}(S(\theta_0))$ or $\bar{r}(S(\theta_1))$, each representing the typical spike train generated by each of the two stimuli. Metric-based distances, $D[\bar{r}(S(\theta_0)), r(f(\theta_0,\theta_1))]$ and $D[\bar{r}(S(\theta_1)), r(f(\theta_0,\theta_1))]$ are calculated using D^{count}, D^{spike}, or $D^{product}$, and the stimulus-template combination, which generates the minimum distance, is used to determine the most likely stimulus.

Although the signal detection theory approach is well established for quantifying the theoretical performance of a neuron (Bair et al. 1994; Newsome et al.

1989), it is typically used on some measure of neuronal output strength, not spike train metrics. Since the spike train templates are unknown, we computed distributions of spike train distances within categories and between categories and computed the optimal performance based on these distributions (see appendix for details). This system is ideal in the sense that it assigns a spike train to the most likely group of spike trains from which it belongs, according to the different metrics. The method is thus useful for comparing spike train metrics to one another using real data. The formal derivation of the theoretical observer's performance, $P_{correct}$, appears in the appendix.

5. *Metric-based data analysis*

The metrics are applied to the data set in the following way. First, the spike trains are numbered such that if there are 2 stimuli, and 8 spike trains recorded in response to each, then the spike trains would be ordered from 1 to 16 (figure 3 - step 2). For each metric parameter, a square matrix of distances between each spike train (i.e. 16×16 elements in this example) is computed, where the value of each element represents the distance between 2 spike trains according to the metric used at the value of the parameter tested. This step is repeated for different values of the metric's free parameter, q or σ. A stimulus-dependent clustering method (Victor and Purpura 1997) is then used to calculate the information entropy, H, in bits (see supplementary materials for details of the entropy calculation). This method does not assume that spike train distances lie in a vector space, and makes no assumptions about parametric relationships between stimulus classes. The information entropy, H, is then plotted against the free parameter of the metric being used (figure 3-step 3). To correct for bias associated with the direct calculation of information (Treves and Panzeri 1995), we used the re-sampling technique in which 10 recalculations of the transmitted information, H, are done following random reassignments of the observed responses to the stimulus classes. The average of these values is used as the estimate of bias due to chance clusters and is subtracted from the information entropy estimate of each neuron (Victor and Purpura 1997). Values of the average bias, under different experimental conditions, are included in the supplementary materials section (figure 1 SM).

The value used for D^{count} is the value at q=0 on the D^{spike} curve (figure 3, step 3, black curve on the left panel). The value used for D^{spike} is the maximum value of

the curve. Similarly the value used for $D^{product}$ is the maximum value on the corresponding $D^{product}$ curve (figure 3, step 3, black curve on the right panel). Error bars were generated using the bootstrap method in which stimulus class indices were randomly selected and information entropy (bias-corrected) and the theoretical observer performance were recomputed. One hundred such recalculations were done for each point, and confidence intervals were computed using the bias corrected and accelerated percentile method (Efron and Tibshirani, 1993). The bootstrap estimates were also used for significance testing for each individual neuron using the Wilcoxon rank-sum test to allow comparisons between the metrics. In experiment 1, for each given stimulus, we recorded 5-12 trials. We analyzed the initial 200 msec window following stimulus onset in 102 neurons from two animals. In experiment 2, for each stimulus, we recorded 4-11 trials, and analyzed the initial 200 msec window following stimulus onset in 103 neurons from the same two animals.

6. Latency analysis

We determined the latency of activity onset in each spike train using a method motivated by Maunsell and Gibson (1992). First, we obtained a set of estimates of the spontaneous or baseline firing rate using the 300 msec pre-trial interval. Since we desired a unique estimate for each spike train, we could not use the PSTH constructed from pooling spike trains. Instead we represented the spike trains using time bins with duration τ. If the estimated firing rate within a time bin was statistically higher than the baseline (using t-test, $p<0.001$), this was determined to be the onset of activity. For small values of τ, this would reduce to the time to the first spike (since, for example, at 1 msec resolution a single spike would yield a firing rate of 1000 Hz). For very large values of τ, temporal resolution limited the ability of the theoretical observer to make decisions. We therefore estimated latencies for a range of values: τ=20, 30, 40, 50, 60, 70, 80, 90, and 100 msec.

We then used standard signal detection theory (Green and Swets, 1966) to compare the performance of a decision maker using firing rate, to one using only the latency values. As with the metric testing, we selected stimuli that elicited maximal performance of the theoretical observer (using the firing rate) subject to the restriction that $0.55<P_{correct}<0.82$. We selected τ for each neuron to maximize the performance of the theoretical observer with respect to the firing rate.

7. *Metric assessment using simulated data*

To better understand the behavior of the two metrics, D^{spike} and $D^{product}$, we applied the metrics to two sets of simulated data with known properties. The first simulation is based on frequency modulation. In this simulation, spike trains are generated from an adaptation of the periodically modulated Poisson distribution (Hunter and Milton 2003), in which the instantaneous firing rate λ(t) is defined as:

$$\lambda(t) = \begin{cases} 0 & \text{if } m\sin(2\pi f_m t) < 0 \\ \lambda_0(1 + m\sin(2\pi f_m t)) & \text{if } 0 \le m\sin(2\pi f_m t) \le 1, \\ 2\lambda_0 & \text{if } m\sin(2\pi f_m t) > 1 \end{cases}$$

where λ_0 is the carrier rate, m is the modulation amplitude and f_m is the modulation frequency. Note that this function is equivalent to a sine function for $0<m<1$, and approaches a square wave as $m \to \infty$. From this distribution, we generated two sets of spike trains representing responses from two theoretical stimuli. The sets of spike trains differed only in the parameter, f_m. We tested this model using two sets of frequency parameters. The low frequency parameters (f_1=5 Hz, f_2=15 Hz) were chosen in part because of the observation of a strong peak in the low frequency range in LFP's recorded from areas MT (Khayat et al. 2008) and LIP (Pesaran et al. 2002), and in part because of the low frequency stable spiking patterns elicited in pyramidal cells *in vitro* (Fellous et al. 2004). We also tested this model at higher frequency parameters (f_1=40 Hz, f_2=50 Hz) motivated by the 30-70 Hz synchronous oscillations found in adjacent neurons in macaque area MT (Kreiter and Singer 1992), and by the 40 Hz peak in the power spectrum of spike trains recorded from area MT (Bair et al. 1994). In these simulations, increasing m makes the task of clustering spike trains to the parent distribution easier.

In order to simulate phase encoding within spike trains (Latham and Lengyel 2008; Montemurro et al. 2008) we repeated the periodically modulated Poisson distribution (PMPD) analysis at low frequency ($f_1=f_2$=15 Hz) and at high frequency ($f_1=f_2$=50 Hz) but with spike trains from each category offset by a phase of Π/4.

The second simulation is based on spike jitter, with no significant difference in the frequency spectra between the two sets of spike trains. In this simulation, two template spike trains are generated from a Poisson process with the same rate parameter, λ. From each template, a set of spike trains is generated by jittering each spike with a Gaussian offset. Formally, let $\{g_{1,\ldots,} g_n\}$ represent event times ~

Poisson(λ). In the first set of spike trains, the j^{th} simulation produces a spike train $g(j)=\{X_{j1},..., X_{jn}\}$, where the element $X_{ji} \sim \text{norm}(g_i, \sigma)$. Thus the spike train *g(j)* is equivalent to *g* (a Poisson spike train), except for a Gaussian jitter applied to each spike. Similarly let $\{h_{1,...,} h_m\}$ represent event times ~Poisson(λ). In the second set of spike trains, the j^{th} simulation produces a spike train $h(j)=\{Y_{j1},..., Y_{jm}\}$, where the element $Y_{ji} \sim \text{norm}(h_i, \sigma)$. σ is the free parameter of the model representing the amount of spike jitter in seconds. For very small values of σ, classification is trivial because all of the spike trains belonging to the *g(j)* group would be very similar to each other, and all of the spike trains belonging to the *h(j)* group would be similar to one another for the same reason. Increasing σ makes the task of clustering spike trains to the parent distribution, $\{g_{1,...,} g_n\}$ or $\{h_{1,...,} h_m\}$, more difficult.

In all cases, 100 simulations were done, each using fifteen trials of 200 msec, with the baseline firing rate, λ_0 =40 Hz. We then applied the three metrics, D^{spike}, $D^{product}$, and D^{count} to these data sets and compared their performance as defined in the previous section. The simulations were conducted over a range of parameters, *m,* in the PMPD simulations and, σ, for the spike jitter simulations.

Results

Metric analysis at the single neuron level

We recorded the responses of each individual neuron to the different direction and contrast combinations. Figure 2*A* shows responses of a typical unit from experiment 1 to the different direction (upper row) and contrast (lower row) combinations of antipreferred and test RDPs. The neuron responds more strongly when the direction of the test RDP was closer to its preferred direction (direction configuration), and when that pattern had higher contrast (contrast configuration). Figure 2*B* shows the responses averaged across trials within the 200 msec period after stimulus onset for both configurations. As with most neurons in our sample, the unit was tuned for changes in the stimulus direction and contrast.

Figure 3 illustrates the metric analysis for another example neuron. The first row shows test RDP stimuli (the antipreferred RDP is not shown in this figure for simplicity) in which the direction of RDPs differs by steps of 15°. This unit gave the largest response to downward motion. The second row shows 200 msec raster plots for each stimulus direction (8 or 9 trials per stimulus in this example). The first step in

the analysis is the selection of the stimuli to be used. We selected two stimuli subject to the restriction that the performance of a theoretical observer at discriminating between them, using a metric based in firing rate alone (D^{count}), was between 55% and 82% (figure 3, step 1). If more than two pairs met this criterion, we chose the pair yielding the highest performance.

This restriction ensured that responses were tuned for spike counts, but that they were not fully discriminable by the theoretical observer when taking into account the response variability. We reasoned that for these stimuli the potential benefit of using timing-based coding schemes (see also Reich et al. 2001b) relative to using spike counts may be more evident than for stimuli eliciting very different firing rates, for which discrimination performance would be close to 100%. We were also motivated by the observation that behavioral thresholds for motion direction discrimination in monkeys are lower than the ones of single neurons in area MT (estimated by signal detection analysis based on spike counts) (Cohen and Newsome, 2009). Spike timing information, in addition to signal pooling (Cohen and Newsome, 2009), may offer a potential solution to this paradox.

Nevertheless, as additional information for the interested reader, we extended our information analysis to all stimulus classes and provide a complete quantification of information entropy using the spike count metric (D^{count}) and the most informative of our time-base metric (D^{spike}) in both experiments (see supplementary material, figure 1 SM).

Following this selection, each of the eight spike trains from one stimulus category is numbered (1 to 8) and each of the spike trains from the other category is numbered (9 to 16). A D^{spike} or $D^{product}$ value is generated between all of the combinations of spike train pairs forming a 16 by 16 matrix for every value of the metric parameter (figure 3, step 2). Two of these matrices are displayed for different parameter values of D^{spike} and $D^{product}$. The metric values are represented by a color map in which blue represents zero, and red represents the maximum value. Theoretically, if spike trains from the same stimulus were all identical, and spike trains from different stimuli were different, this matrix would appear blue in the upper left quadrant and lower right quadrant, and red in the lower left quadrant and upper right quadrant. Each matrix is then collapsed into a single information entropy value (figure 3, step 3 - black line), and a single value of the probability of correct response, $P_{correct}$ (figure 3, step 3, red line), as described in the appendix.

Continuing with this example neuron, an examination of the entropy curve using the D^{spike} metric reveals that at q=0, the entropy estimate is 0.21 bits. As the cost of shifting a spike, q, is increased, the function achieves a maximum value of 1.0, at q=32, but beyond q=64, the information entropy decreases precipitously. Thus for this neuron the D^{spike} metric at q=32 sec^{-1} could potentially transmit about five times more information than D^{count}. Similarly, the right panel demonstrates that the information content increases for $D^{product}$ as σ is increased from 2^{-11} sec (approximately 0.5 msec) to 2^{-7} sec (approximately 8 msec) from 0.54 to 1.0. Both of these increases were significant (Wilcoxon rank-sum, $p<0.0001$). A similar trend occurs analyzing the $P_{correct}$ function (red curve using the right-sided y-axis). Analyzing for D^{spike}, at q=0, $P_{correct}$=0.71, and as q increases to 32 sec^{-1}, $P_{correct}$ increases to a maximum value of 0.87. Thus, the theoretical observer would be able to perform about 16% better at correctly assigning a spike train to its category using spike timing as with the D^{spike} metric. A similar pattern is seen with $D^{product}$, although the maximum value (at $\sigma= 2^{-7}$ sec) is 0.78, somewhat lower than for D^{spike} (at q=32 sec^{-1}). Nevertheless, both D^{spike} and $D^{product}$ yielded significant increases in performance (Wilcoxon rank-sum, $p<0.0001$).

Metric analysis at the population level - Experiment 1

Figure 4 shows mean entropy data pooled across all neurons. We applied the analysis illustrated in figure 3 to each unit and then averaged across cells. In both animals, there was significantly higher information entropy of stimulus clustering for both contrast (A) and direction (B), when using the D^{spike} metric compared to the D^{count} metric (2-tail paired t-test, $p<0.0001$ in both cases). On the single cell level, 54 out of 102 of the cells yielded significant increases (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method (see methods section)).

In both animals, M1 (black bars) and M2 (grey bars), the mean entropy of contrast stimulus clustering was 0.26 and 0.31 bits using the D^{count} metric, 0.57 and 0.63 bits using the $D^{product}$ metric, and 0.65 and 0.70 bits using the D^{spike} metric (figure 4A). The differences using $D^{product}$ versus D^{count} were significant in both animals (2-tail paired t-test, $p<0.0001$), with 53 of the neurons exhibiting significant increase (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method), however in 16 cells there was significantly higher information entropy using D^{count} than using $D^{product}$. The differences using $D^{product}$ and D^{spike} over the population of cells were not significant

(2-tail paired t-test, p=0.11, and p=0.33). On the single cell level, 44 out of 102 neurons yielded significantly higher entropy using D^{spike} over $D^{product}$, compared to 25 cells in which entropy was higher using $D^{product}$ (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method).

In the direction trials (figure 4B), the mean entropy of stimulus clustering was 0.24 and 0.23 bits using the D^{count} metric, 0.48 and 0.63 bits using the $D^{product}$ metric, and 0.59 and 0.65 bits using the D^{spike} metric. In both animals, pairwise comparisons between D^{spike} and $D^{product}$ were not significant (2-tail paired t-test, p=0.08, p=0.16), but between $D^{product}$ and D^{count} were significant (2-tail paired t-test, p=0.00027, $p<0.0001$). 58 cells exhibited significantly increased information using D^{spike} compared to D^{count} (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method). 60 neurons exhibited significantly increased information using $D^{product}$ compared to D^{count}, however, D^{count} produced increased estimates of information in 12 neurons. Comparing D^{spike} and $D^{product}$, 39 cells yielded higher entropy using D^{spike}, and 23 cells yielded higher entropy using $D^{product}$ (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method).

We applied signal detection analysis to the spike train data to quantify the performance of a theoretical observer using an ideal discrimination strategy subject to the three choices of metrics (see methods). Figure 5A summarizes the results of contrast discrimination with bar heights representing mean, and error bars standard error. In M1, the mean probability of a correct decision using D^{count} was $P_{correct}= 0.72$, using $D^{product}$ was $P_{correct}= 0.77$, and using D^{spike} was $P_{correct}= 0.81$. In M2, these values were $P_{correct}=0.71$ for D^{count}, $P_{correct}= 0.78$ for $D^{product}$, and $P_{correct}= 0.82$ for D^{spike}. In both animals, there were significant differences between the performance using D^{count}, and D^{spike} (2-tail paired t-test, $p<0.001$), but not between the performance of $D^{product}$ and D^{spike} (2-tail paired t-test, p=0.079, p=0.16). In both monkeys, performance using $D^{product}$ was significantly better than using D^{count} (2-tail paired t-test, p=0.0068, and p=0.05).

Figure 5B and C demonstrate this trend across individual neurons. They plot the performance of each neuron using D^{count} on the x-axis, and D^{spike} (B) or $D^{product}$ (C) on the y-axis. Figure 5B demonstrates a clear trend of improved performance for almost every cell using D^{spike} over D^{count} for contrast discrimination. This increase was statistically significant in 71 cells (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method). Figure 5C shows a different trend, with several neurons' spike

trains yielding better performance with D^{count} than with $D^{product}$. Of these, 36 were significant. On the other hand, 55 neurons yielding improved performance using $D^{product}$. Figure 5D replots the performance of each neuron using $D^{product}$ on the x-axis, and D^{spike} on the y-axis. Although there is significant variability between neurons, there is a trend of improved performance with the D^{spike} metric over $D^{product}$. Performance was significantly better using D^{spike} over $D^{product}$ in 52 cells, compared to 37 cells in which performance was significantly better using $D^{product}$ (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method). The parameters used for the $D^{product}$ (figure 5E) and D^{spike} (figure 5F) metrics ranged between neurons. In the case of $D^{product}$, σ ranged from 0.2 msec (2^{-12} sec) to 250 msec (2^{-2} sec), with a mean value of 42 msec. This distribution is skewed, however, with its peak occurring at ~1 msec (2^{-10} sec). In the case of D^{spike}, the parameter q ranged from 0 sec^{-1} to 128 sec^{-1} with a population mean of 26.1 sec^{-1}. This corresponds to the majority of the distribution occurring at q=16 sec^{-1}, and q=32 sec^{-1}. Of note, 16 cells (out of 102) did not exhibit any improved performance for values of q>0 sec^{-1} suggesting that in this subset of cells there is no role for temporal precision coding.

We repeated this analysis for direction discrimination (figure 6). For the two monkeys the mean performance values were $P_{correct}$=0.70 and $P_{correct}$=0.71 using D^{count}, $P_{correct}$=0.71 and $P_{correct}$=0.74 using $D^{product}$, and $P_{correct}$=0.78 and $P_{correct}$=0.80 using D^{spike} (figure 6A). For both animals there was significant pairwise improvement from D^{count} to D^{spike} (2-tail paired t-test, $p<0.0001$). In both cases there was no significant difference using $D^{product}$ versus D^{count} (2-tail paired t-test , p=0.61, p=0.33). Figures 6B and C plot joint probabilities of each neuron's performance (as in figure 5). Figure 6B demonstrates improved performance for the majority of neurons using D^{spike} compared to D^{count}. This improvement was statistically significant in 66 cells (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method). Similar to the contrast tuning case, this trend does not exist for $D^{product}$, with several cells yielding worse performance with $D^{product}$ than with D^{count} (40 cells had significant improvement using $D^{product}$, whereas 41 had significantly worse performance using $D^{product}$). Again, in general D^{spike} seems to perform the best.

The latter trend is further highlighted by reanalyzing the performance of each neuron using $D^{product}$ on the x-axis, and D^{spike} on the y-axis (figure 6D), demonstrating a clear trend of improved performance using D^{spike} over $D^{product}$. This improvement was significant in 60 cells compared to 13 cells in which performance was better

using $D^{product}$ (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method). Analysis of the metric parameters used in the direction discrimination task (figure 6E and F) yielded similar results as in the contrast discrimination task. In the case of $D^{product}$, σ ranged from 0.2 msec to 250 msec, with a mean value of 45 msec. This distribution peak occurred at approximately 4 msec (2^{-8} sec). In the case of D^{spike}, the parameter q ranged from 0 sec^{-1} to 64 sec^{-1} with a population mean of 18 sec^{-1}, and the distribution peak occurring at q=32 sec^{-1}. 18 cells did not exhibit any improved performance for values of q>0 sec^{-1}.

The mean values for the cost parameter (18.0 sec^{-1} in the contrast discrimination case and 26.1 sec^{-1} in the direction discrimination case) suggests integration of temporal information occurring approximately over 40 to 60 msec. The analysis of the smoothing parameter, σ, in the $D^{product}$ metric is more difficult to interpret. Although the mean values were 42 and 45 msec, the distribution was very skewed toward much shorter estimates. There are two potential explanations for this skew. First, as figure 5C shows, several cells performed substantially worse using $D^{product}$ than D^{count}, and therefore the $D^{product}$ parameter may be biologically meaningless in these cases. Second, as in the single neuron case shown in figure 3 (bottom right), several neurons had a plateau in performance over a range of σ in the $D^{product}$ metric suggesting that the actual parameter value may not be very informative.

ASSESSING METRICS' BEHAVIOR USING SURROGATE DATA SETS: The calculation of D^{spike} used in figures 3, 4, and 5 involves using the cost parameter, q. In order to determine if any potential bias was imposed by the extra parameter, we recomputed D^{count} and D^{spike} on a surrogate data set constructed by replacing each spike train with an artificial spike train generated from a Poisson process with a rate parameter equal to the firing rate of the original spike train. Since this surrogate data set should contain no temporal information other than chance clustering, any improvement in performance using D^{spike} over D^{count} would be due to sample bias, and this could be used to estimate the bias in the original data. This analysis was applied to the contrast configuration (figure 7A), and to the direction configuration (figure 7B) using the rate parameters corresponding to the data from M1 and M2 separately (different colors). In the contrast configuration (figure 7A), this computation yielded mean performance values of $P_{correct}$=0.658 using D^{count} and $P_{correct}$=0.663 using D^{spike} in M1. This difference was statistically significant (2-tail paired t-test, p=0.03), although the mean

difference was very small, representing 6% of the main effect shown in figure 6A. In M2, the values were $P_{correct}$=0.657 using D^{count} and $P_{correct}$=0.667 using D^{spike}. This difference was not significant (2-tail paired t-test, p=0.23). In the direction configuration (figure 7B), the Poisson recalculation yielded values of $P_{correct}$=0.647 using D^{count} and $P_{correct}$=0.653 using D^{spike} in M1, and $P_{correct}$=0.640 using D^{count} and $P_{correct}$=0.643 using D^{spike} in M2. These differences were not significant (2-tail paired t-test, p=0.2 and p=0.5). Thus the majority of the main effect of improved performance using D^{spike} compared to D^{count} is not attributable to bias arising from the added parameter. Furthermore, the difference between D^{spike} and $D^{product}$ cannot be explained by differences in the number of free parameters, since both make use of one free parameter relating to the timescale of temporal integration.

Temporal correlation between spikes within a spike train have been previously proposed as a means of robust signaling, overcoming synaptic noise (Salinas and Sejnowski 2001). In order to assess the contribution that temporally correlated spikes within a spike train might have on temporal coding, we generated a second set of surrogate data for reanalysis. In this analysis, for every spike train, each spike time was randomly reassigned to another spike train *within* the same stimulus category. This method maintains the spike count for every spike train (and consequently D^{count} is unaffected), and maintains spike times for every category, but eliminates the effect of temporal correlation between spikes. Figures 7C and D plot the results of this analysis for contrast and direction discrimination. The difference in population means between original and resampled data was small (~0.015) and not significant (2-tail paired t-test, contrast: p=0.18, p=0.51; direction: p=0.31, p=0.46). Therefore, the contribution of temporally correlated spikes does not appear to be the main source of usable spike timing. Furthermore, since this recalculation effectively destroys the original sequences of interspike intervals, potential information stored in these intervals is also unlikely to be used.

EFFECT OF THE ANALYZED RESPONSE PERIOD ON THE METRICS' BEHAVIOR: All of the preceding analyses have been done using the first 200 msec of response following stimulus presentation. It may be that this has an influence on the contribution of spike timing to direction and contrast coding and that such a contribution becomes smaller during later periods of the response. In order to test this hypothesis, we reapplied the analysis to later response periods also using a 200 msec interval, but starting 350 msec

after the stimulus onset (figure 8A for contrast conditions, figure 8B for direction conditions). The trend of improved performance using D^{spike} over D^{count} exists over this time interval (2-tail paired t-test, contrast: p=0.005, p=0.0016, direction: p=0.0026, p=0.005), though this effect was significantly smaller than the one using the initial 200 msec. This occurred both because of improved performance using D^{count} and reduced performance using D^{spike}. The clearest difference between analyses over this interval and over the initial 200 msec, was the substantial and significant reduction in performance using the $D^{product}$ metric. This occurred in the data from both monkeys, and in both conditions (2-tail paired t-test, contrast: $p<0.0001$, p=0.0030, direction: $p<0.0001$, p=0.013).

In summary, our choice of analysis time window does have an effect on the behavior of the different metrics, the performance of D^{count} improved, the one of D^{spike} slightly decreased and the one of $D^{product}$ suffered a major drop. The fact that the differences between D^{spike} over D^{count} were much larger than the ones isolated in figure 7A and B using the surrogate data sets suggests that the superior performance of the former metric was not due to the additional free parameter. Thus, although somewhat reduced relative to early response periods, spike timing (quantified by D^{spike}) seems to provide a coding advantage during the later sustained response period.

Population results – Experiment 2

In order to account for any possible effect of the stimulus configuration in experiment 1, we repeated the analysis of information entropy of stimulus clustering and theoretical observer performance on a second data set. Here, only one RDP moving in different directions was presented inside the cell's RF (see experiment 2 methods). We concentrated on the metrics D^{count} and D^{spike} because they yielded the largest and most consistent difference in the previous analysis (see supplementary materials and figure 2 SM for results using $D^{product}$). Figure 9 plots the result of this analysis. Using the D^{count} metric, the mean entropy values in M1 and M2 were 0.40 and 0.25, respectively (figure 9A). When D^{spike} was used, the mean entropy values increased to 0.72 and 0.83. In both animals, the differences were significant (2-tail paired t-test: $p<0.001$ in both cases, with significant increases in 45 cells using Wilcoxon rank-sum, $p<0.05$ with the bootstrap method).

Figure 9B plots the theoretical observer model performance in both animals' data sets using the metrics D^{spike} and D^{count}. Using D^{count}, mean performance values were $P_{correct} = 0.70$ and $P_{correct} = 0.72$. Using D^{spike} improved performance to $P_{correct} =$

0.81 and $P_{correct}$ = 0.85. In both animals, the differences were significant (2-tail paired t-test: $p<0.001$ in both cases with significant increases in 56 cells using Wilcoxon rank-sum, $p<0.05$ with the bootstrap method). Figure 9C plots the joint probability histograms for performance using D^{count} versus D^{spike} for each neuron. For most neurons, there was a clear improvement in performance using D^{spike}. Figure 9D plots the values of the cost parameters that were used in these estimates. The majority of neurons had their ideal performance optimized at cost values of q=8 sec^{-1} and q=16 sec^{-1}. Thus, the results of the metric analysis using data from experiment 2 are similar to the results of experiment 1. The aspects of spike timing quantified by D^{spike} provide a coding advantage for both, direction and contrast.

Analysis of response latencies

The previous results lead to the question of which features of spike timing quantified by D^{spike} is responsible for improved performance. One potential candidate is the latency of activity onset (Raiguel et al. 1999). We reasoned that since latency relies on the time of the first spike or spike burst, the D^{spike} metric undoubtedly captures such a feature. For example, two spike trains in which the first spike or burst occurs at the same time will lead to a low value of D^{spike} relative to spike trains in which they occur at different times since the only difference would be the cost of shifting these initial spikes in time. Therefore, one would anticipate that if latency was the feature that made D^{spike} the best performing metric, then comparing the performance using latency against the one using firing rate should be similar to comparing D^{spike} vs. D^{count}.

Figure 10 summarizes the results of our latency analysis. In this case we have used a more classical approach than our metric analysis, i.e., standard signal detection analysis (see methods). This allows us to interpret our results within the context of previous studies. However, the methods of signal detection theory are in principle similar to the ones using the metric-based analysis. The panels in the first row of this figure (A-C) plot the mean firing rate (across several trials) in the stimulus condition which produced the smaller number of spikes (non-preferred) against the mean firing in the condition which produced the larger number of spikes (preferred condition) (see methods for details of selection). As expected, all of these values fall below the unity line, demonstrating that mean firing rate was consistently lower in the non-preferred condition.

The panels in the second row of figure 10 (D-F) plot the mean latency of activity onset (across several trials) in the non-preferred condition against the mean latency of activity onset in the preferred condition. The vast majority of points in all three cases fall above the unity line suggesting that our estimate of mean latency was consistently lower in the preferred relative to the non-preferred condition.

However, these mean firing rate and latency measurements do not provide an estimate of these parameters variability in individual trials. Since we were interested in a measurement of performance that would incorporate such variability, we subjected these two measurements to separate signal detection analyses and compared their results.

In all three cases (figure 10G-I), for both monkeys, the theoretical observer performed substantially better in individual trials using firing rate than using the best performing measurement of latency of activity onset. This trend was also seen at the single neuron level with significantly better performance using firing rate over latency in 72 and 74 cells (out of 102) in the contrast and direction configurations of experiment 1, compared to only 10 and 8 cells with significantly better performance using latency over firing rate (Wilcoxon rank-sum, $p<0.05$). In experiment 2, 58 cells performed significantly better using firing rate than using latency of activity onset, compared to 17 cells in which the opposite was true (Wilcoxon rank-sum, $p<0.05$).

In summary, the results of this analysis suggest that latency alone cannot be the temporal feature providing benefit to the D^{spike} metric over D^{count}. In fact, firing rate seems to be more informative than latency. Interestingly, the D^{spike} metric is sensitive to both of these spike train features, firing rate and latency.

Assessing metrics behavior using simulated spike trains

In order to further explore some properties of the spike trains that would yield the previous results, spike trains were generated as described in the methods section, and the performance of the theoretical observer was computed. Figure 11 plots the results of 100 simulations (in each subplot), each with 15 trials of 200 msec generated for each category. The mean performance is plotted with error bars representing standard errors. Note that in figures 11A-D the error bars are smaller than the data points.

Figure 11A shows the performance (see methods) at distinguishing spike trains generated by the periodically modulated Poisson distribution (PMPD) model with

frequency parameter f_1=5 Hz from those generated by the PMPD model with frequency parameter f_2=15 Hz. No phase offset was used in this simulation. The D^{spike} and $D^{product}$ metric performed much better than D^{count} over the entire tested range of modulation amplitude, *m*. Over this range, the $D^{product}$ outperformed the D^{spike} metric. When PMPD spike trains were generated with higher frequencies, f_1=40 Hz and f_2=50 Hz the relationship between metric performance was the same. Over the range of modulation amplitudes, $D^{product}$ yielded better performance than D^{spike}. To assess how phase offset might affect each metric, we repeated the simulation with the same carrier frequency (15 Hz in figure 11C, 50 Hz in figure 9D) in both categories, but with a phase offset of Π/4. In both cases, the relationship between the metrics' performance was again the same. Thus, for these period data and over the range of modulation amplitudes, $D^{product}$ yielded better performance than D^{spike}.

We finally used temporal jitter of individual spikes as a parameter (instead of modulation amplitude). In this scenario, the metric performance was very different (figure 11E). Over the tested range of the spike jitter (σ), D^{spike} performed significantly better than $D^{product}$, while D^{count} performed generally the worst, although the performance of these latter metrics seems to overlap over a range of spike jitters. These results resemble the ones described in the different data figures of both experiments.

In summary, the results of these analyses indicate that periodicity within the spike train (better accounted for by $D^{product}$) was not the feature that caused the superior performance of D^{spike} in our data sets. On the other hand, they suggest that our spike trains from single neurons more closely resemble the simulated data with spike jitter, containing non-periodic time information.

Discussion

In this study, we compared the information entropy derived from spike trains produced by MT neurons in response to RDPs with varying contrasts and directions using the metrics D^{count}, D^{spike} and $D^{product}$. In contrast to previous studies, we concentrated our analysis on stimuli in which a theoretical observer model, using spike counts (D^{count}), would perform suboptimally (below 82%) at discriminating between stimuli with different contrast and direction. We derived the observer's performance using the three different metrics.

Both, the information and signal detection theoretic approaches have been previously used to quantify the stimulus-encoding capabilities of visual cortical neurons (Bair et al. 1994; Newsome et al. 1989; Reich et al. 2001b; Samonds and Bonds 2004), and provide complementary insights into neural coding. While the information entropy approach provides a measure of stimulus clustering that applies to a metric space, and is independent of any assumed geometry of responses to stimuli, the signal detection (theoretical observer model) approach treats spike train distances as vectors within a Euclidean space and quantifies the performance of a downstream (at a later stage) decision-maker.

In area MT, the metric space approach has not been previously reported. Applying this approach and the information theoretic analysis to our MT data resulted in a substantial increase in information content (over twice as much in both conditions and in both monkeys) using either the D^{spike} or the $D^{product}$ metric compared to using D^{count}. The theoretical observer analysis suggests that this increase in information only reliably occurs using D^{spike} and could translate to an almost 10% improvement in discrimination performance (in the range of 55% to 82%).

Both D^{spike} and $D^{product}$ measure similarity of temporal structure between spike trains. Overall, however, the $D^{product}$ metric failed to provide reliable improvements of performance compared to D^{count}. To better understand why performance improved using the D^{spike} metric, and less so (if at all) using $D^{product}$, we repeated our analysis on two types of spike trains generated by the PMPD and spike jitter models, with known temporal structure. In the PMPD model, temporal frequency or temporal phase is used as a coding parameter. In the spike jitter model, random (Poisson) spike trains with no predominant oscillatory structure are identical within categories except for a uniformly distributed temporal jitter.

The potential for superiority of $D^{product}$ at detecting differences in frequency and phase is intuitive since it is a vector space computation and the deterministic components of the PMPD model for different values of f_0 are orthogonal. Similarly, a metric such as D^{spike}, in which cost is associated with spike shifting, ought to perform best for spike trains differing by random jitter of random spike times. This is demonstrated in figure 11E. D^{spike} appears to perform better than $D^{product}$ when temporal spike coding exists and is not based on changes in periodicity. Conversely, $D^{product}$ tended to perform better than D^{spike} when spike coding is based on changes in periodicity. From this perspective, we suggest that the overall superior performance of

D^{spike} over $D^{product}$ in classifying the spike trains from our recordings implies that the temporal structure of area MT neuronal spike trains, at least in our experiments, is not the result of intrinsic subthreshold oscillations within the neuron or reverberant network activity that produce oscillatory behavior. This interpretation is also consistent with the large interburst intervals and the 40 Hz peak in the power spectra, not seen in the Fourier spectra of MT neuronal discharges reported by Bair et al. (1994), which they interpreted as a lack of pacemaker activity in MT neurons.

The reliability of spike timing in the responses of MT neurons has been previously investigated (Bair and Koch 1996; Buracas et al. 1998; Fellous et al. 2004; Osborne et al. 2004; Masse and Cook 2008). In the first of these studies, spike jitter standard deviation was computed between spike trains with identical stimuli (RDPs with varying degrees of coherency). The highest temporal precision (lowest spike jitter standard deviation) was evoked by low coherence RDPs, and was practically absent for 100% coherence patterns, like the ones we used in this study. However, we did find temporal structure in the spike trains elicited by such patterns.

We consider at least three likely reasons for the differences between our results and those of the previously mentioned study. The first is that the metric space approach using the D^{spike} metric allows for the detection of more subtle temporal coding schemes because it allows for spike deletions and insertions at varying costs related to moving a spike in time. Secondly, Bair and Koch (1996) analyzed the sustained response (until 2 seconds of stimulus presentation) - precisely where we found the least effect of temporal precision (see Figure 8). Consistent with this hypothesis, Reich and colleagues (Reich et al. 2001b) also reported the role of spike timing in contrast coding by V1 neurons mainly during response onset. Thirdly, these authors used stimuli moving in the neurons' preferred direction and speed, whereas our analysis was restricted to stimuli that elicited responses lying close together along the neurons' direction and contrast response functions, where the benefit of spike time based coding may be more evident.

It should also be pointed out that the stimuli in experiment 1 of our study had two RDPs, one that changed contrast or direction, and one that always had the same contrast, and always moved in the neuron's antipreferred direction. This may result in a decrease in the overall amount of spikes (e.g., *via* random deletions), thus decreasing any effect of precise spike timing. Nevertheless, under these conditions our theoretical observer analysis demonstrated significant potential benefit in using a

temporal precision code. It is possible that our scenario better reflects real environmental conditions. For example, moving stimuli in many situations are adjacent to one another, and under these circumstances, an observer must make accurate perceptual judgments. In addition, the similar results that we obtained in experiment 2, in which only one RDP was inside the neuron's RF, suggests that our results can be generalized to many situations.

Fellous et al. (2004) reported evidence for temporal structure in the firing pattern of area MT neurons. They also used the normalized inner product of the spike density function (Schreiber 2003) as their measure of similarity. However, they used a vector based clustering algorithm and therefore required the additional step of histogram-reshaping with a sigmoid function. The fact that they found substantially improved performance using this measure, whereas the improvement we found seem to be smaller, may be a consequence of the stimulus used. In that study, the animal viewed moving Gabor patches that may have elicited a higher oscillatory response by stimulus-locking due to the spatiotemporal frequency component of the patch. If that were the case, as our PMPD simulations demonstrated, we would expect a high performance from a metric based on an inner product such as $D^{product}$.

In agreement with our results, Osborne and colleagues (Osborne et al. 2004) analyzed the information content of direction selective responses of MT neurons (using firing rate) as a function of time. They found that the first 100 msec of the spike train after the response latency (~80 msec) contains 80% of the information about the stimulus direction. Our results suggest that even more stimulus information can be gleaned within this time interval by making use of temporal precision. Additionally, there is psychophysical evidence that temporal integration of motion direction by humans, in tasks using highly coherent RDPs, does not dramatically change after 250 msec of stimulus presentation (Watamaniuk and Sekuler 1992).

In order to demonstrate the potential advantage of temporal coding during perceptual task, we introduced a theoretical observer model in which the network downstream to the recorded MT neuron "reads" its output (the spike train), compares to a template, and "solves" a discrimination task. Potential physiological correlates for neuronal representation of a time-varying signal (the template) may be coincidence detection of spatially correlated spikes, variations in the weighing of synaptic decays within a dendritic field, or different configurations of time varying membrane voltage

channels (Di Maio 2008; Mockett and Hulme 2008). These possibilities, however, are highly speculative and remain to be experimentally tested.

One may argue that the improved performance of the D^{spike} metric over D^{count} in the contrast discrimination case is trivial since decreasing stimulus contrast demands increased temporal integration for reliable detection. However, the effect of delayed neuronal activity affects the D^{count} metric as much as D^{spike} since the overall number of spikes in a 200 msec window would decrease with delayed onset. Further, this effect could not account for the differences seen in the direction discrimination case. The fact that potential improvements in coding for both of these stimulus parameters could be achieved using spike timing, suggests that this effect may be a feature of the input to area MT rather than originating in area MT itself since contrast discrimination may occur in areas upstream from MT. This agrees with the findings of temporal coding of contrast in V1 of anesthetized monkeys (Reich et al. 2001a).

Finally, one may ask whether a specific temporal feature of the spike train such as latency to the first spike or burst of activity could explain the advantage seen with D^{spike} relative to the other metrics. The data shown in figure 10 demonstrate that this alone is not the case. Furthermore, it suggests that the timing of the spikes after the first burst of activity, at least until 200ms from stimulus onset, also contributes to the advantage observed by using the D^{spike} metric.

One very likely explanation for our results is that more than one feature of spike trains – such as a hybrid of spike count, latency, and possibly the shape of the transient "envelope" (to which $D^{product}$ would be sensitive) may have produced the increase in information and performance seen with the D^{spike} metric. Thus, any method that would examine one of these individual features in isolation will necessarily underperformed D^{spike}, and underestimate how the brain uses different types of available information.

In conclusion, this study used a metric based analysis to demonstrate a significant *potential* advantage in the use of spike time coding by area MT neurons when discriminating stimuli that vary in direction and contrast, but that are not easily discriminable by a theoretical observer using spike counts. Although we cannot, under these experimental circumstances demonstrate that the brain *is* using spike times to boost discrimination performance, it is likely that the brain uses a source of easily available information. If so, it may explain some differences in performance between single neurons in monkeys - using rate codes as a means to quantify information in

neurometric functions - compared with the organisms behavior (Newsome et al. 1989; Snowden and Braddick 1990; Snowden et al. 1991; Cohen and Newsome 2009). Future studies must better establish cause-effect relationships between spike time information and the neural representations of stimulus attributes that underlie our perceptual skills.

Appendix

The performance of a theoretical observer subject to a metric-based strategy is calculated here. This observer is ideal in the sense that it makes the best possible decision subject to a metric based strategy. First, let us consider a general mathematical form of the stimulus,

$$f_i(\theta_0,\theta_1) = G_i \cdot \theta_1 + (1 - G_i)\cdot \theta_0,$$

where G_i is a random variable with binomial distribution assuming a value of 0 or 1 with equal probability, and i is the trial number. The neuron's response can be described by the response function $r_i(f(\theta_0,\theta_1))$. In this model, the system uses template spike trains, $\bar{r}(S(\theta_0))$, and $\bar{r}(S(\theta_1))$. We define the decision that the system makes in terms of a decision function,

where D^{metric} is the metric being tested, and may represent D^{count}, D^{spike}, $D^{product}$ or any

$$dec_{D^{metric}}(f(\theta_0,\theta_1)) = \begin{cases} 0, & \text{if } P[G=0|f(\theta_0,\theta_1)] > P[G=1|f(\theta_0,\theta_1)]; \\ 1, & \text{if } P[G=1|f(\theta_0,\theta_1)] > P[G=0|f(\theta_0,\theta_1)]; \end{cases}$$

other spike train metric. To derive the probability of this decision maker being correct, $P_{correct}$, we first derive the probability of an incorrect response, $P_{\text{incorrect}}$:

$$\begin{aligned} P_{incorrect} &= P\left(dec_{D^{metric}}\left(f(\theta_0,\theta_1)\right) \neq G\right) \\ &= P(G=1)\cdot P\left[dec_{D^{metric}}\left(f(\theta_0,\theta_1)\right) = 0 | G=1\right] + P(G=0)\cdot P\left[dec_{D^{metric}}\left(f(\theta_0,\theta_1)\right) = 1 | G=0\right] \\ &= P(G=1)\cdot P\left[D^{metric}\left\{g_i\left(f(\theta_0,\theta_1)\right),\bar{g}\left(S(\theta_1)\right)\right\} > D^{metric}\left\{g_i\left(f(\theta_0,\theta_1)\right),\bar{g}\left(S(\theta_0)\right)\right\} | G=1\right] + \\ & P(G=0)\cdot P\left[D^{metric}\left\{g_i\left(f(\theta_0,\theta_1)\right),\bar{g}\left(S(\theta_1)\right)\right\} \leq D^{metric}\left\{g_i\left(f(\theta_0,\theta_1)\right),\bar{g}\left(S(\theta_0)\right)\right\} | G=0\right]. \end{aligned}$$

To simplify the derivation, let us introduce the following notation. Let $d_{r_i,0} = D^{metric}\{g_i(f(\theta_0,\theta_1)),\bar{g}(S(\theta_0))\}$, that is, the distance using a given metric between the real spike train and the template spike train for the stimulus $S(\theta_0)$, for each trial, i. Similarly, let $d_{r_i,1} = d^{metric}\{g_i(f(\theta_0,\theta_1),\bar{g}(\theta_1)\}$. We then define the set of all such distances with the first template, $D_{r,0} = \{d_{r_1,0},d_{r_2,0},...,d_{r_n,0}\}$ and with the second template, $D_{r,1} = \{d_{r_1,0},d_{r_2,0},...,d_{r_n,0}\}$. Also assume that the stimuli are equiprobable, that is P(G=1)=P(G=0)=0.5. Then for the i^{th} trial,

$$P_{i,incorrect} = 0.5 \cdot \left\{ P\left[d_{r_i,0} < d_{r_i,1} \mid G = 1 \right] + P\left[d_{r_i,0} > d_{r_i,1} \mid G = 0 \right] \right\}$$
$$= 0.5 \cdot \left\{ P\left[d_{r_i,0} - d_{r_i,1} < 0 \mid G = 1 \right] + P\left[d_{r_i,0} - d_{r_i,1} > 0 \mid G = 0 \right] \right\}.$$

The distances in the above expression, $d_{r_i,0}$ and $d_{r_i,1}$, cannot be computed directly since the templates, $\bar{r}(\theta_0)$ and $\bar{r}(\theta_1)$, are unknown. However, we can estimate the above expression by using each spike train as a template and calculating cross metrics between spike trains. To do this we define the set of distances:

$$\hat{d}_{11} = \left\{ D^{metric}\left[r_1, r_2\right], D^{metric}\left[r_1, r_3\right], \ldots, D^{metric}\left[r_1, r_m\right], D^{metric}\left[r_2, r_3\right], \ldots, D^{metric}\left[r_2, r_m\right], \ldots, D^{metric}\left[r_{m-1}, r_m\right] \right\},$$

where *m* is the number of trials with stimulus $S(\theta_0)$. Note that this set has $\alpha = \sum_{i=2}^{m} \sum_{j=1}^{i-1} 1$ elements. Similarly,

$$\hat{d}_{10} = \left\{ D^{metric}\left[r_{m+1}, r_1\right], D^{metric}\left[r_{m+1}, r_2\right], \ldots, D^{metric}\left[r_{m+1}, r_m\right], \ldots, D^{metric}\left[r_{m+n}, r_1\right], \ldots, D^{metric}\left[r_{m+n}, r_m\right] \right\},$$

where *n* is the number of trials with stimulus $S(\theta_1)$. This set has $\beta = mn$ elements. Finally,

$$\hat{d}_{00} = \left\{ D^{metric}\left[r_{m+2}, r_{m+1}\right], D^{metric}\left[r_{m+3}, r_{m+1}\right], \ldots, D^{metric}\left[r_{m+3}, r_{m+2}\right], \ldots, D^{metric}\left[r_{m+4}, r_{m+1}\right], \ldots, \right.$$
$$\left. D^{metric}\left[r_{m+4}, r_{m+3}\right], \ldots, D^{metric}\left[r_{m+n}, r_{m+1}\right], \ldots, D^{metric}\left[r_{m+n}, r_{m+n-1}\right] \right\},$$

which contains $\gamma = \sum_{i=m+2}^{m+n} \sum_{j=m+1}^{i-1} 1$ elements. The estimate of $P_{incorrect}$ based on the recorded spike trains is:

$$\hat{P}_{incorrect} = \frac{1}{2} \cdot \left[\frac{1}{\beta\gamma} \sum_{i=1}^{\beta} \sum_{j=1}^{\gamma} \begin{cases} 1, \text{ if } \hat{d}_{i,j}10(i) - \hat{d}_{i,j}11(j) < 0 \\ 0, \text{ if } \hat{d}_{i,j}10(i) - \hat{d}_{i,j}11(j) \geq 0 \end{cases} + \frac{1}{\alpha\beta} \sum_{i=1}^{\alpha} \sum_{j=1}^{\beta} \begin{cases} 1, \text{ if } \hat{d}_{i,j}10(i) - \hat{d}_{i,j}00(j) < 0 \\ 0, \text{ if } \hat{d}_{i,j}10(i) - \hat{d}_{i,j}00(j) \geq 0 \end{cases} \right].$$

Finally, the probability of a correct response of the theoretical observer subject to a metric-based analysis is $\hat{P}_{correct} = 1 - \hat{P}_{incorrect}$. We refer to this entity throughout the text as $P_{correct}$.

Acknowledgements: We thank Walter Kucharski for his technical assistance. We thank Dr. Eric Cook, and Dr. Nick Masse for their useful comments on the manuscript.

Grants: This research was supported by grants from the CFI, CRC, CIHR, and The Ottawa Hospital Department of Surgery, and The Ottawa Hospital Division of Neurosurgery.

References

Bair W, and Koch C. Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comput* 8: 1185-1202, 1996.

Bair W, Koch C, Newsome WT, and Britten KH. Power spectrum analysis of bursting cells in area MT in the behaving monkey. *J Neurosci* 14: 2870-2892, 1994.

Barlow H, and Tripathy SP. Correspondence noise and signal pooling in the detection of coherent visual motion. *J Neurosci* 17: 7954-7966, 1997.

Born RT, and Bradley DC. Structure and function of visual area MT. *Annu Rev Neurosci* 28: 157-189, 2005.

Buracas GT, Zador AM, DeWeese MR, and Albright TD. Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron* 20: 959-969, 1998.

Carandini M, Heeger DJ, and Movshon JA. Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17: 8621-8644, 1997.

Cohen MR, and Newsome WT. Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J Neurosci* 29: 6635-6648, 2009.

Di Lorenzo PM, and Victor JD. Taste response variability and temporal coding in the nucleus of the solitary tract of the rat. *J Neurophysiol* 90: 1418-1431, 2003.

Di Maio V. Regulation of information passing by synaptic transmission: A short review. *Brain Res* 1225: 26-38, 2008.

Dubner R, and Zeki SM. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res* 35: 528-532, 1971.

Efron B, and Tibshirani RJ. *An Introduction to the Bootstrap.* Boca Raton, FL: Chapman and Hall/CRC, 1998.

Fellous JM, Houweling AR, Modi RH, Rao RP, Tiesinga PH, and Sejnowski TJ. Frequency dependence of spike timing reliability in cortical pyramidal cells and interneurons. *J Neurophysiol* 85: 1782-1787, 2001.

Fellous JM, Tiesinga PH, Thomas PJ, and Sejnowski TJ. Discovering spike patterns in neuronal responses. *J Neurosci* 24: 2989-3001, 2004.

Goldberg DH, Victor JD, and Gardner D. The spike train analysis toolkit: A component of a computational neuroinformatic resource. In: *Society for Neuroscience*. Washington, DC: 2005.

Green DM, and Luce MW. Parallel psychometric functions from a set of independent detectors. *Psychol Rev* 82: 483-486, 1975.

Green DM, and Swets JA. *Signal detection and recognition by human observers.* New York, NY: John Wiley and Sons, 1966.

Holt GR, Softky WR, Koch C, and Douglas RJ. Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *J Neurophysiol* 75: 1806-1814, 1996.

Hu D, Chelvanayagam DK, and Vidyasagar TR. Irregularity in cortical firing in vivo cannot be explained by inhibition. *Proceedings of the Australian Neuroscience society* 13: 2002.

Hubel DH, and Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195: 215-243, 1968.

Hunter JD, and Milton JG. Amplitude and frequency dependence of spike timing: implications for dynamic regulation. *J Neurophysiol* 90: 387-394, 2003.

Hunter JD, Milton JG, Thomas PJ, and Cowan JD. Resonance effect for neural spike time reliability. *J Neurophysiol* 80: 1427-1438, 1998.

Khayat P, Niebergall R, and Martinez-Trujillo J. Attentional modulation of local field potentials response tuning to direction and contrast in macaque's area MT. In: *Society for Neuroscience*. Washington, DC: 2008.

Kreiter AK, and Singer W. Oscillatory Neuronal Responses in the Visual Cortex of the Awake Macaque Monkey. *Eur J Neurosci* 4: 369-375, 1992.

Kruse W, and Eckhorn R. Inhibition of sustained gamma oscillations (35-80 Hz) by fast transient responses in cat visual cortex. *Proc Natl Acad Sci USA* 93: 6112-6117, 1996.

Latham PE, and Lengyel M. Phase coding: spikes get a boost from local fields. *Curr Biol* 18: R349-351, 2008.

Liu J, and Newsome WT. Local Field Potential in Cortical Area MT: Stimulus Tuning and Behavioral Correlations. *J Neurosci* 26: 7779-7790, 2006.

Malach R, Schirman TD, Harel M, Tootell RB, and Malonek D. Organization of intrinsic connections in owl monkey area MT. *Cereb Cortex* 7: 386-393, 1997.

Martinez-Trujillo J, and Treue S. Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron* 35: 365-370, 2002.

Martinez-Trujillo JC, and Treue S. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol* 14: 744-751, 2004.

Masse NY, and Cook EP. The effect of middle temporal spike phase on sensory encoding and correlates with behavior during a motion-detection task. *J Neurosci* 28: 1343-1355, 2008.

Maunsell JH, and Gibson JR. Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol* 68: 1332-1344, 1992.

Mockett BG, and Hulme SR. Metaplasticity: new insights through electrophysiological investigations. *J Integr Neurosci* 7: 315-336, 2008.

Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, and Panzeri S. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr Biol* 18: 375-380, 2008.

Moulden B, Kingdom F, and Gatley LF. The standard deviation of luminance as a metric for contrast in random-dot images. *Perception* 19: 79-101, 1990.

Movshon JA, Thompson ID, and Tolhurst DJ. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J Physiol* 283: 53-77, 1978.

Newsome WT, Britten KH, and Movshon JA. Neuronal correlates of a perceptual decision. *Nature* 341: 52-54, 1989.

Niebergall R, Khayat P, and Martinez-Trujillo J. Neural correlates of multiple spotlights of attention in area MT during multiple-object tracking. In: *Society for Neuroscience*. Washington, DC: 2008.

Osborne LC, Bialek W, and Lisberger SG. Time course of information about motion direction in visual area MT of macaque monkeys. *J Neurosci* 24: 3210-3222, 2004.

Pesaran B, Pezaris J, Sahani M, Mitra P, and Andersen R. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat Neurosci* 5: 805-811, 2002.

Raiguel SE, Xiao D-K, Marcar VL, and Orban GA. Response latency of macaque area MT/V5 neurons and its relationship to stimulus parameters. *J Neurophysiol* 82: 1944-1956, 1999.

Reich DS, Mechler F, and Victor JD. Formal and attribute-specific information in primary visual cortex. *J Neurophysiol* 85: 305-318, 2001a.

Reich DS, Mechler F, and Victor JD. Temporal coding of contrast in primary visual cortex: when, what, and why. *J Neurophysiol* 85: 1039-1050, 2001b.

Rodieck RW, Kiang NY, and Gerstein GL. Some quantitative methods for the study of spontaneous activity of single neurons. *Biophys J* 2: 351-368, 1962.

Salinas E, and Sejnowski TJ. Correlated neuronal activity and the flow of neural information. *Nat Rev Neurosci* 2: 539-550, 2001.

Samonds JM, and Bonds AB. From another angle: Differences in cortical coding between fine and coarse discrimination of orientation. *J Neurophysiol* 91: 1193-1202, 2004.

Schreiber S. A new correlation-based measure of spike timing reliability. *Neurocomputing* 7, 2003.

Shadlen MN, and Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18: 3870-3896, 1998.

Snowden RJ, and Braddick OJ. Differences in the processing of short-range apparent motion at small and large displacements. *Vision Res* 30: 1211-1222, 1990.

Snowden RJ, Treue S, Erickson RG, and Andersen RA. The response of area MT and V1 neurons to transparent motion. *J Neurosci* 11: 2768-2785, 1991.

Softky WR, and Koch C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci* 13: 334-350, 1993.

Treue S, and Martinez-Trujillo JC. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399: 575-579, 1999.

Treves A, and Panzeri S. The upward bias in measures of information derived from limited data samples. *Neural Comput* 7: 399-407, 1995.

Victor JD, and Purpura KP. Metric-space analysis of spike trains: theory, algorithms and application. *Network* 8: 127-164, 1997.

Watamaniuk SN, and Sekuler R. Temporal and spatial integration in dynamic random-dot stimuli. *Vision Res* 32: 2341-2347, 1992.

Watson AB. Gain, noise, and contrast sensitivity of linear visual neurons. *Vis Neurosci* 4: 147-157, 1990.

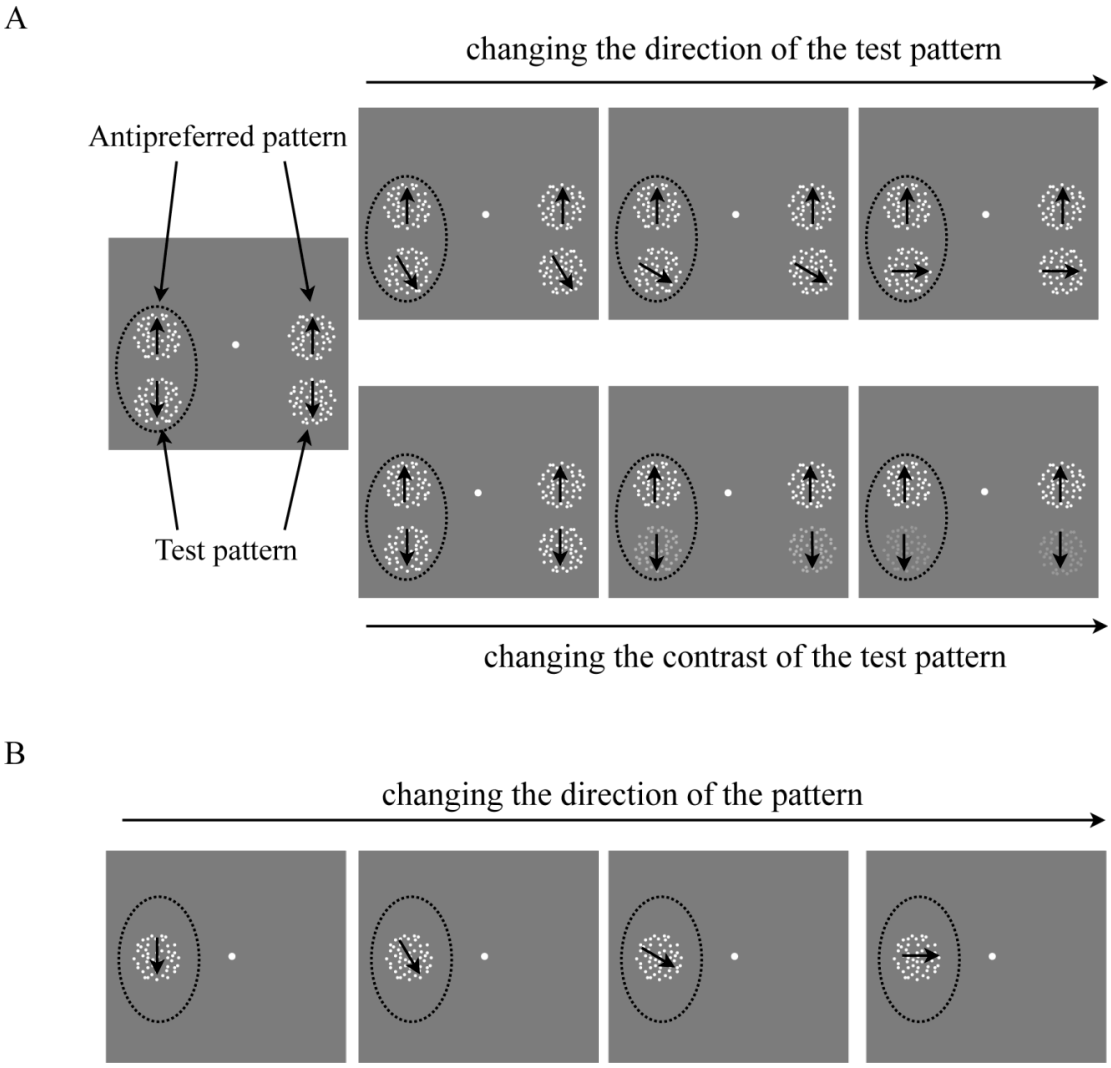


Figure 1. Stimulus configuration. In experiment 1 (A), a test and a null (antipreferred) RDP were positioned inside the RF (dashed circle) and a similar pair was positioned in the opposite hemifield. In the direction configuration, the direction of motion of the test RDP was varied (top). In the contrast configuration, the contrast of the test RDP was varied (bottom). In experiment 2 (B), only one RDP was positioned inside the neuron's RF. The direction of motion was altered between conditions. In both cases the monkey detected a contrast change in the fixation dot (central spot).

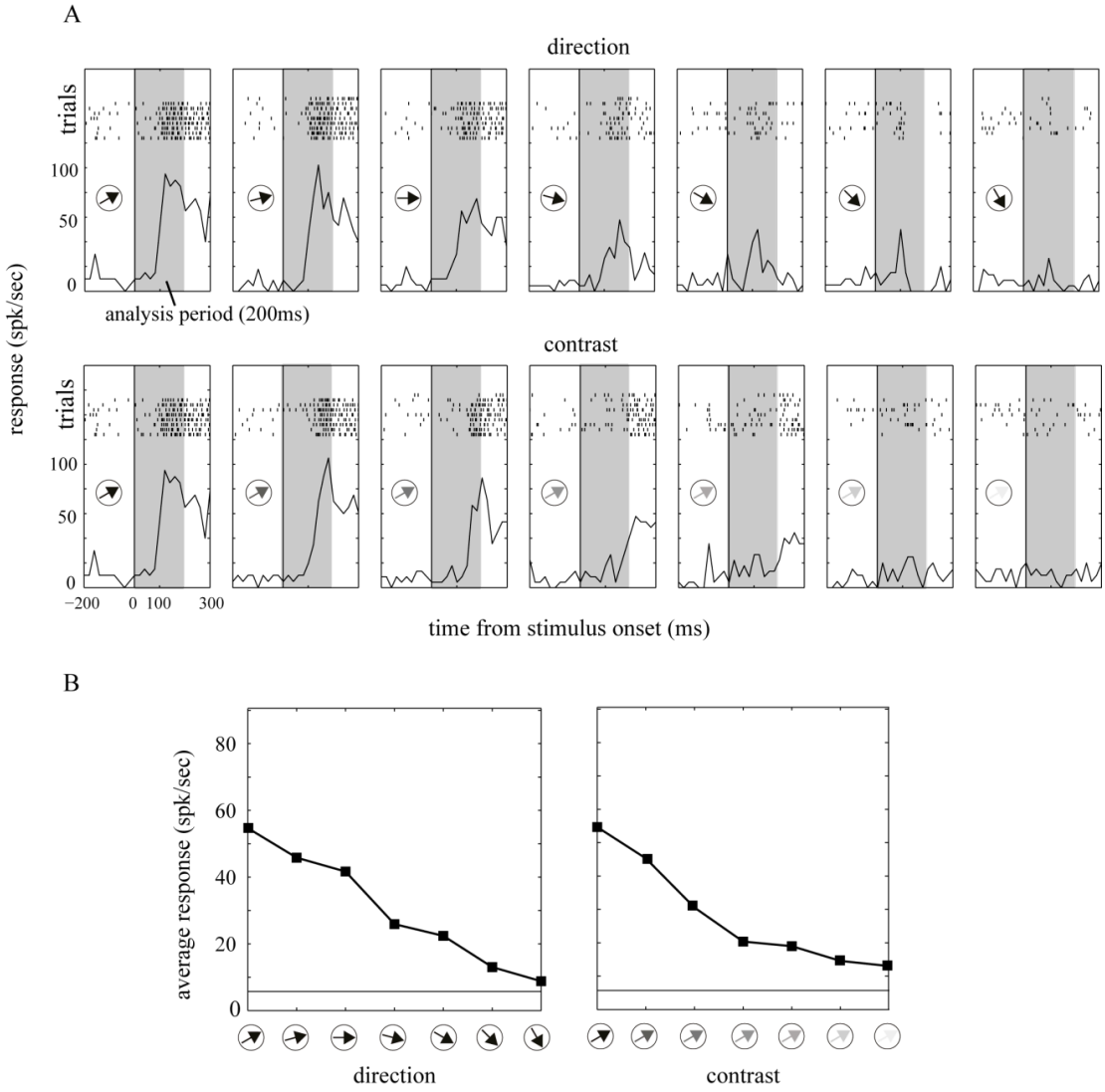


Figure 2. (A) Example neuron response to different combinations of the contrast and direction configurations of the first experiment. In each panel, the abscissa displays time from stimulus onset and the ordinate the response in spikes per second. The upper part of each panel displays raster plots of different trials used to generate the peristimulus time histograms (PSTHs). The gray shaded area represents the response time period used during the analyses. A bin width of 20 msec was used to generate the PSTHs. (B) Mean responses across the test pattern direction (left panel) or contrast (right panel). The neuron's preferred motion direction was 60° clockwise from vertical up (see arrows).

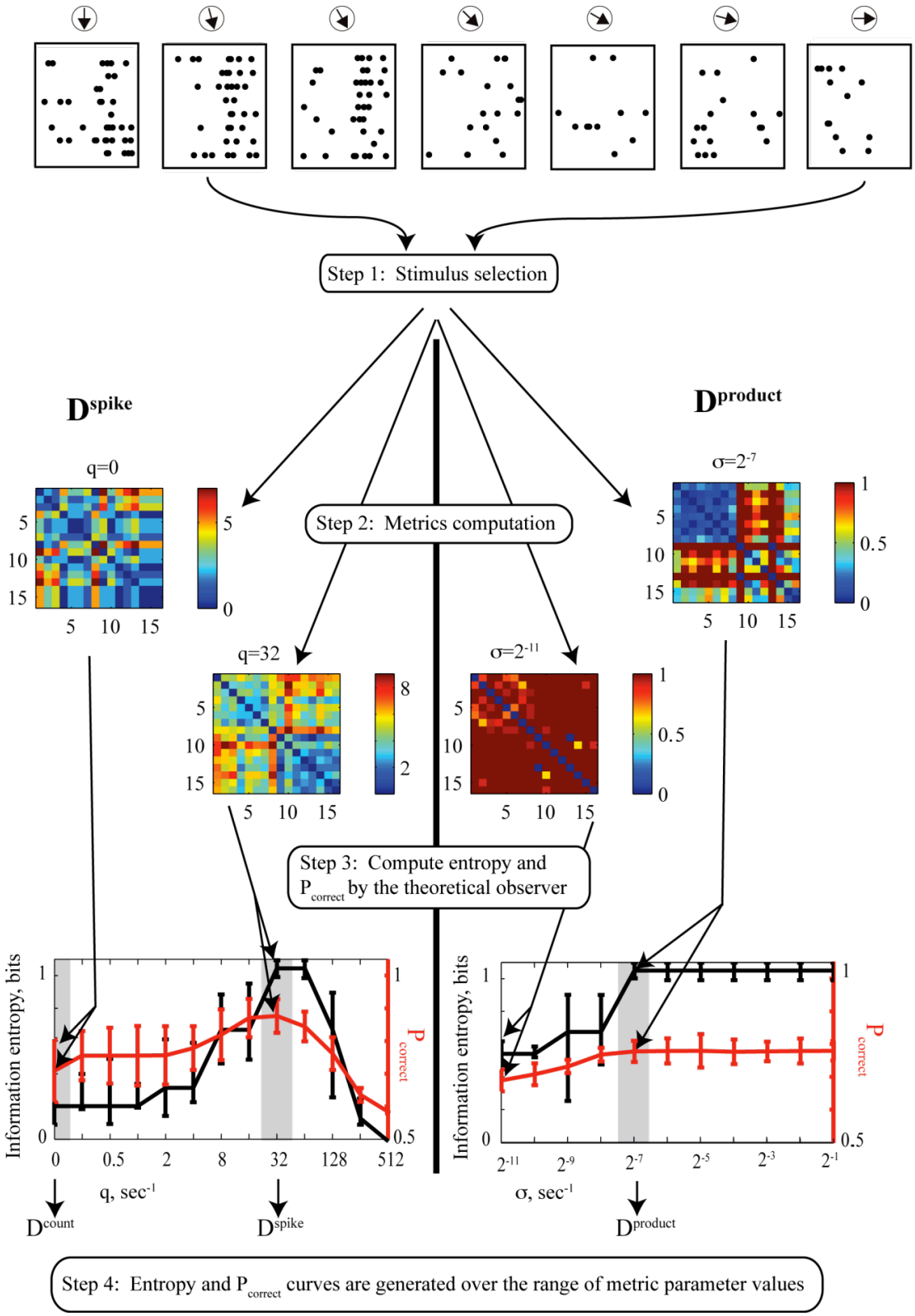


Figure 3. Metric and theoretical observer analysis in a single neuron example. Steps are delineated from top to bottom. First, multiple spike trains are recorded in response to different test pattern directions. Step 1 of the analysis is to select two sets of stimuli that evoke maximal $P_{correct}$ between 0.55 and 0.82 using the D^{count} metric. The

corresponding sets of spike trains are then used to compute matrices (step 2-see methods for details) of D^{spike} values (left) and $D^{product}$ values (right) for different values of the cost parameter (left) or Gaussian kernel parameter (right). Each of these matrices is then summarized as a point (step 3) on the information curve (black) and a point in the $P_{correct}$ curve (red). Finally, the points q=0 on both D^{spike} curves are taken as the D^{count} values, and the maximum values are taken as the values for D^{spike} or $D^{product}$ (step 4). Error bars represent 95% confidence intervals as estimated using the bootstrap method (see methods).

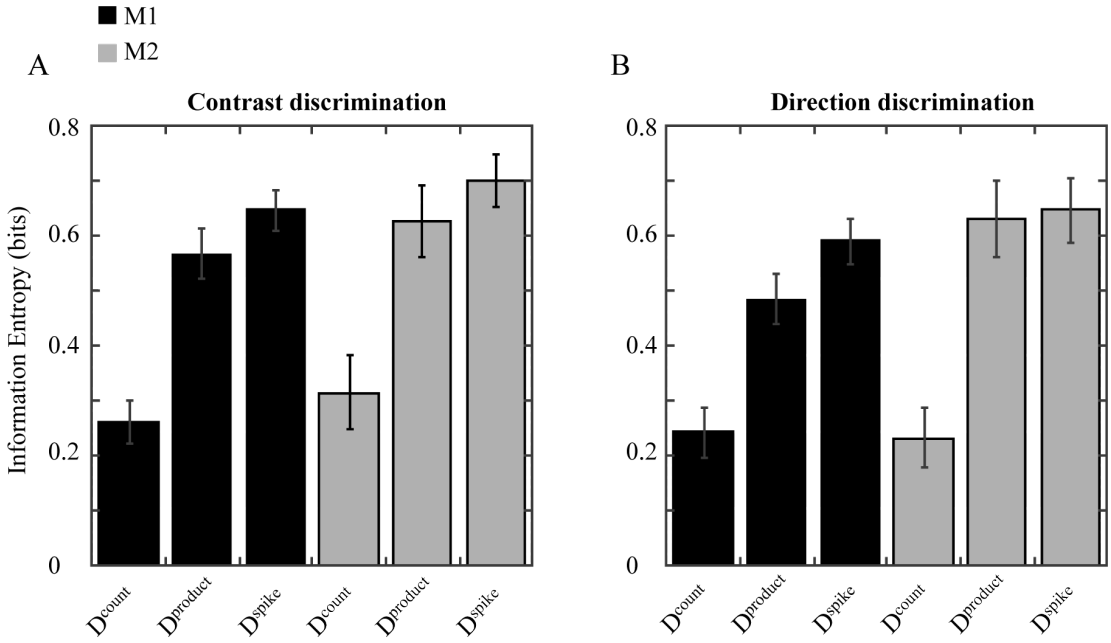


Figure 4. Information entropy of stimulus clustering using different metrics on spike trains recorded from area MT neurons responding to one of two contrasts of the test stimulus at the optimal direction (A), and one of two test stimulus directions at high contrast (B). In both cases the antipreferred stimulus was always accompanying the test stimulus (see figure 1A). Error bars represent standard errors. Black and grey bars represent data from M1 and M2 respectively.

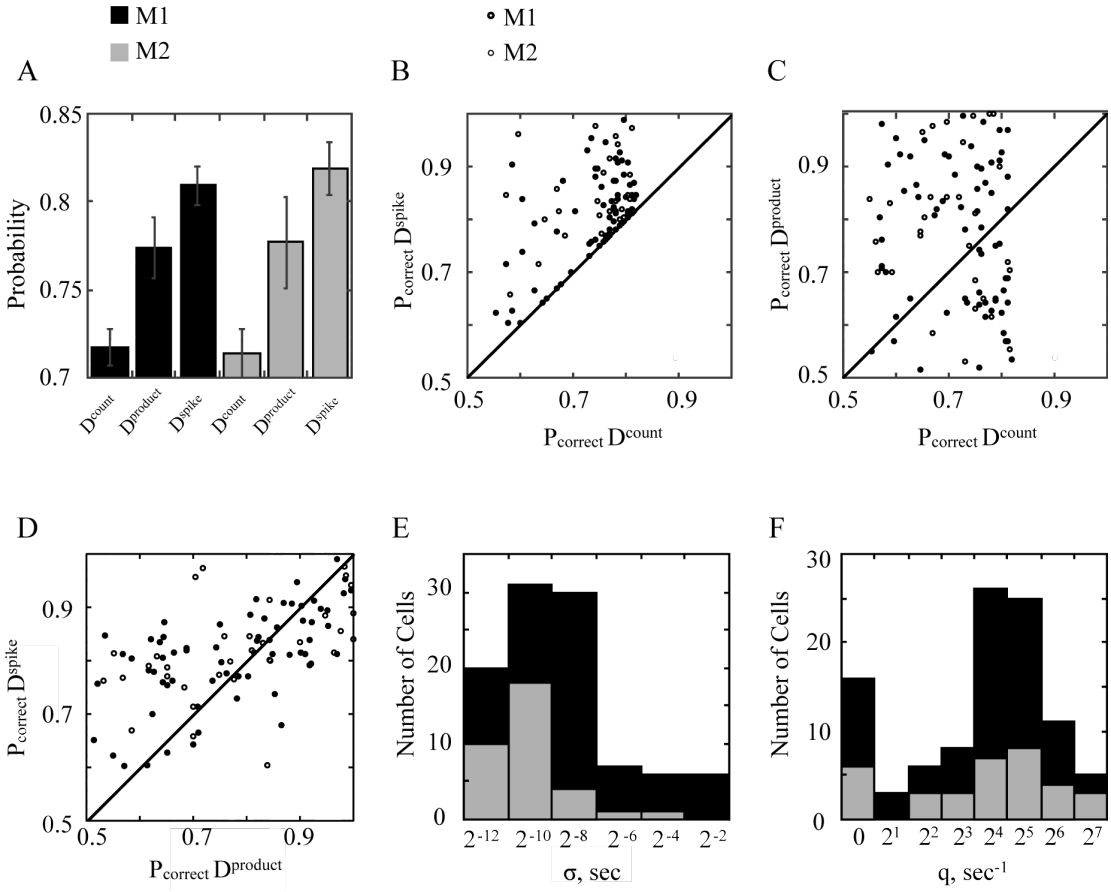


Figure 5. Performance of the theoretical observer model in a contrast discrimination task using D^{count}, $D^{product}$, and D^{spike} metrics (A). Joint probability graphs plotting performance of each neuron using a D^{spike} versus D^{count} (B), $D^{product}$ versus D^{count} (C), and D^{spike} versus $D^{product}$ (D). The solid line in B, C, and D is the line of equivalency. Histograms of the model parameters are shown in E for the $D^{product}$ metric, and in F for the D^{spike} metric.

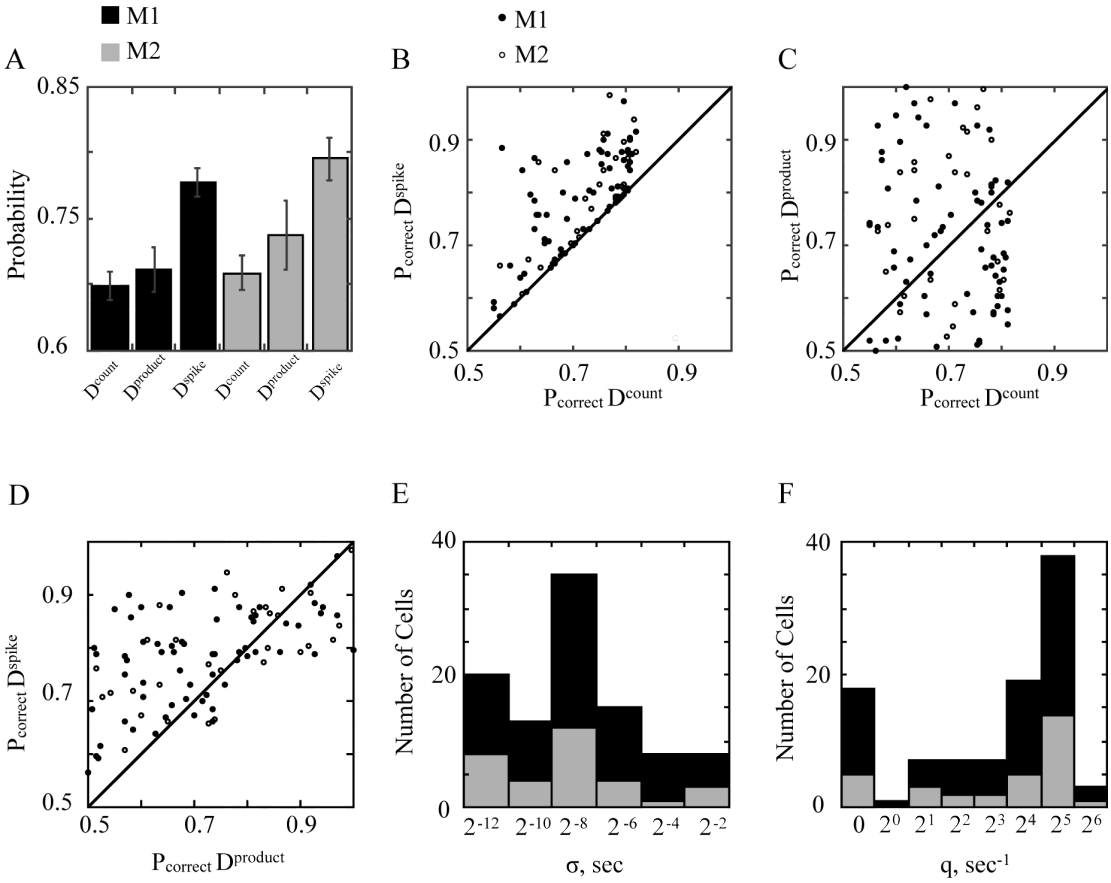


Figure 6. Performance of the theoretical observer model in a motion direction discrimination task using D^{count}, $D^{product}$, and D^{spike} metrics (A). Joint probability graphs plotting performance of each neuron using D^{spike} versus D^{count} (B), $D^{product}$ versus D^{count} (C), and D^{spike} versus $D^{product}$ (D). The solid line in B, C, and D is the line of equivalency. Histograms of the model parameters are shown in E for the $D^{product}$ metric, and in F for the D^{spike} metric.

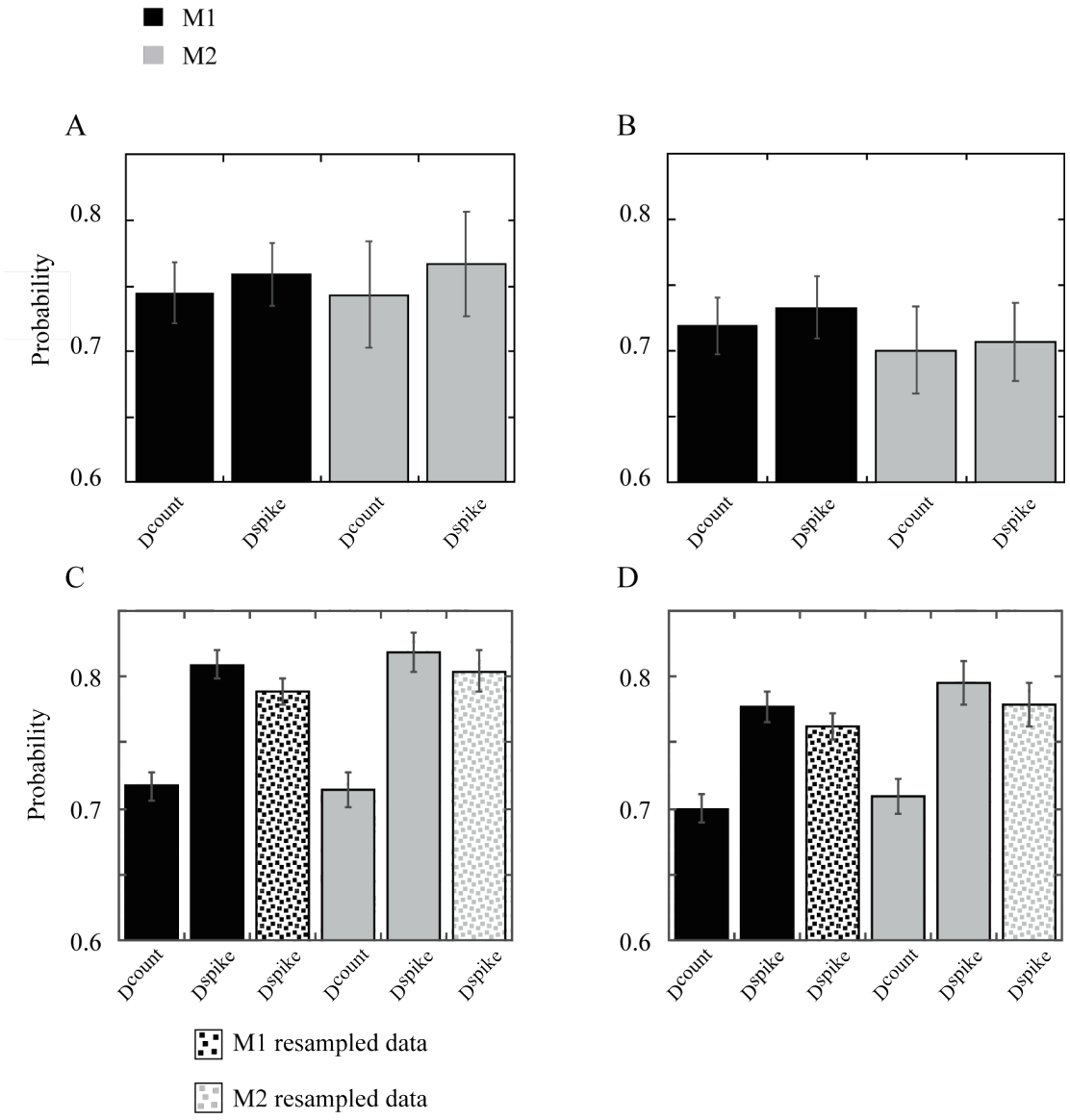


Figure 7. Theoretical observer performance using D^{spike} and D^{count} on surrogate data sets constructed by replacing the spike train generated from each trial (from experiment 1) with a Poisson process in the contrast categories (A) and direction categories (B). Theoretical observer performance using D^{spike} and D^{count} on real (as in figure 4 and 5) and surrogate data sets constructed by randomly exchanging spike times between trials of the same category in contrast conditions (C) and direction conditions (D).

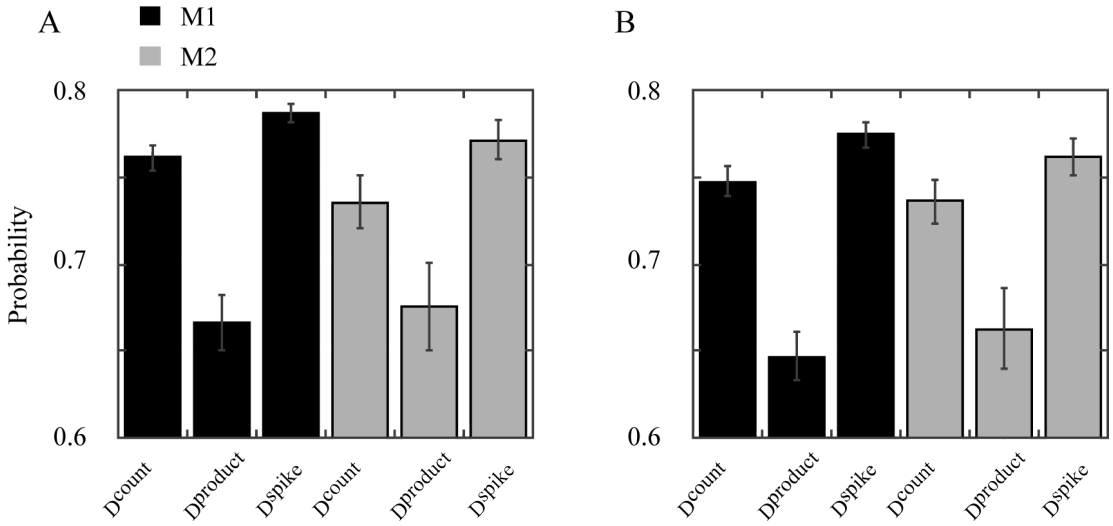


Figure 8. Theoretical observer performance using sustained response of MT neurons to the stimulus (350-550 msec after stimulus presentation) for contrast discrimination (A) and motion direction discrimination (B).

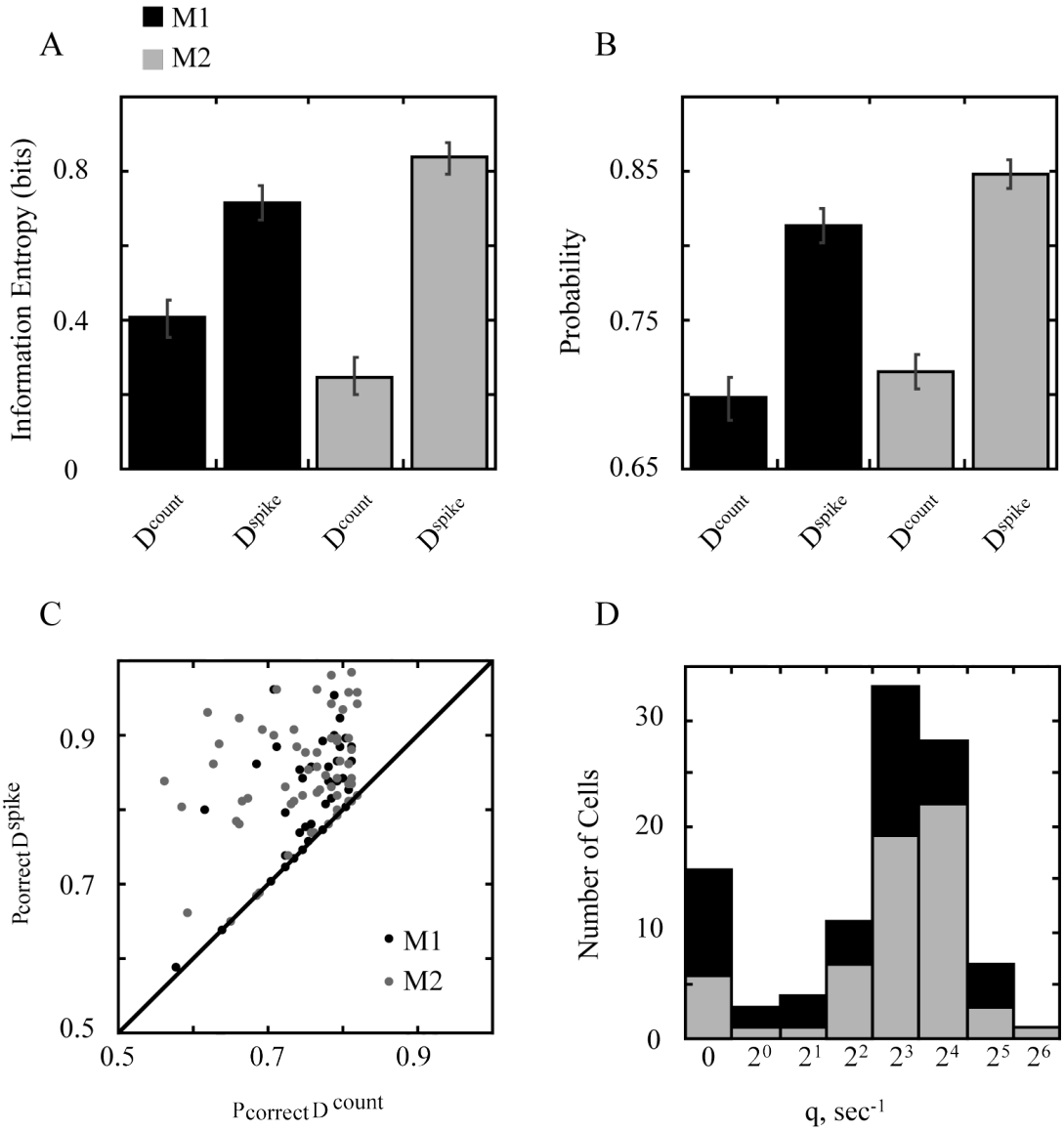


Figure 9. Results of experiment 2. (A) Information entropy of stimulus clustering using the D^{count}, and D^{spike} metrics. (B) Theoretical observer performance in a direction discrimination task using the D^{count}, and D^{spike} metrics. (C) Joint probability graphs plotting performance of each neuron using a D^{spike} versus D^{count}. (D) Histogram of the cost parameters used with the D^{spike} metric.

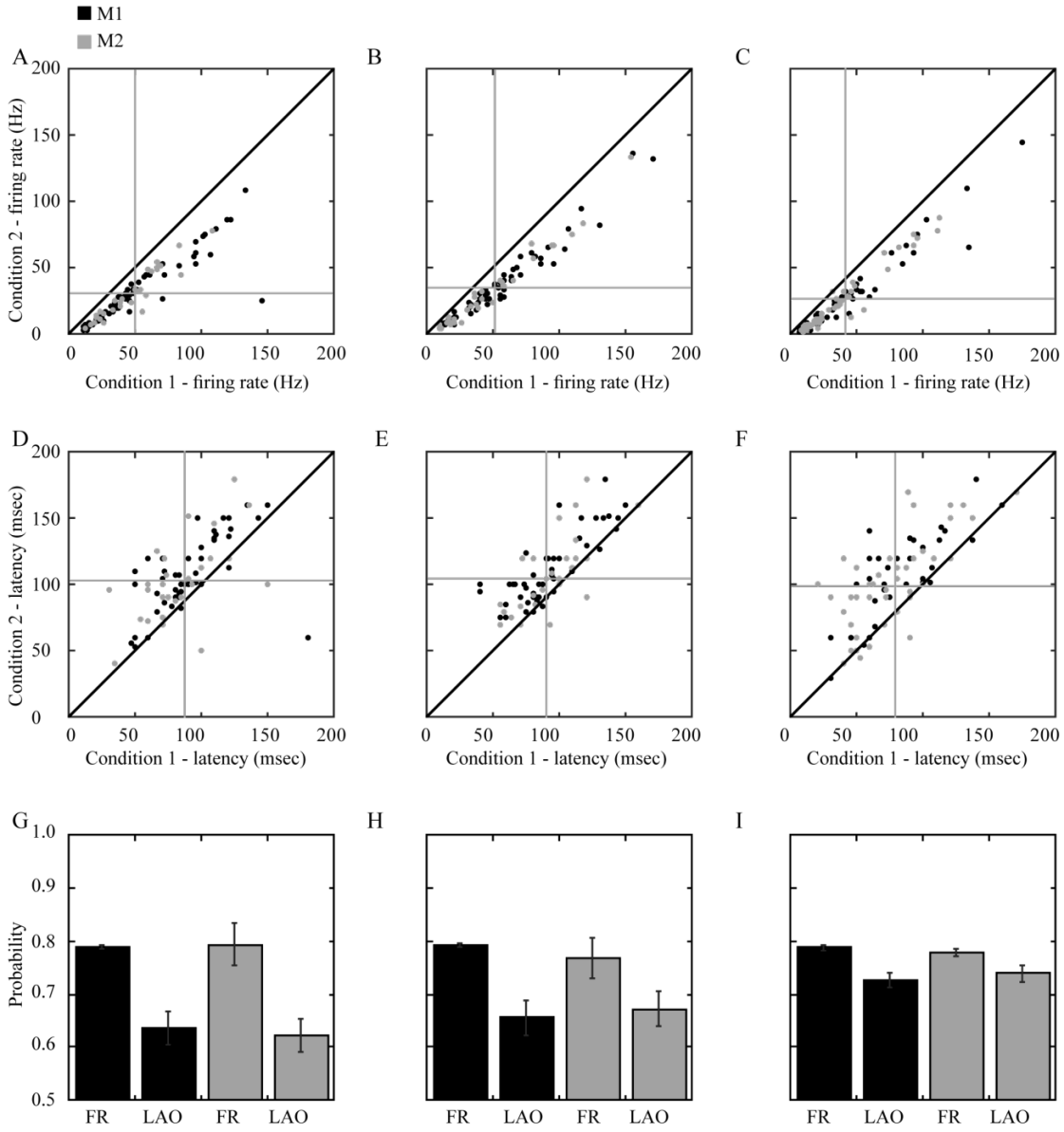


Figure 10. Theoretical observer performance using either a firing rate measure (FR) or the latency of activity onset (LAO). A-C: mean firing rate for each neuron at condition 1 and condition 2 (see results) in contrast (A) and direction (B) configuration of experiment 1, and in experiment 2 (C). D-F plot the mean latency for each neuron at condition 1 and condition 2 in contrast (D) and direction (E) configuration of experiment 1, and in experiment 2 (F). In A-F, the grey horizontal and vertical bars represent the mean values across neurons. Theoretical performance in the contrast condition of experiment 1 (G), the direction configuration of experiment 1 (H), and in experiment 2 (I).

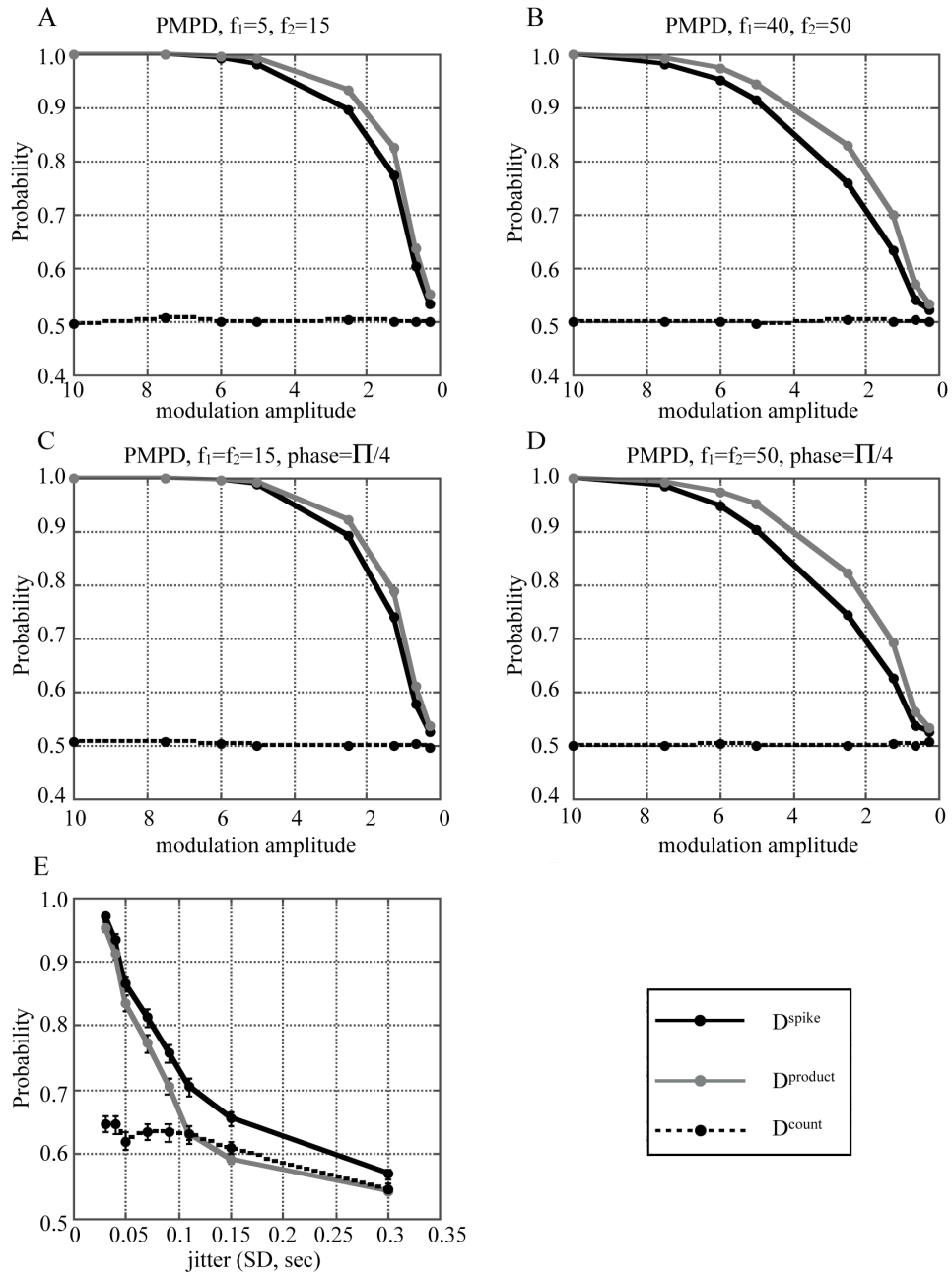


Figure 11. Theoretical observer performance using the D^{spike}, $D^{product}$, and D^{count} metrics for simulated spike trains generated by the PMPD model with zero phase-offset and carrier frequency parameters f_1= 5 Hz, f_2=15 Hz (A), carrier frequency parameters f_1= 40 Hz, f_2=50 Hz (B), between-category phase-offset of Π/4 and carrier frequency f_1=f_2=15 Hz (C), between-category phase-offset of Π/4 and carrier frequency f_1=f_2=50 Hz (D), and the spatial jitter model (E).

Supplementary Material

SM 1. Information Entropy of Stimulus Clustering

We used the method of Victor and Purpura (1997) to quantify the information entropy using spike train metrics. We review the details of this method here. Suppose that N_{tot} spike trains were elicited from one of the stimulus classes $s_1,s_2,\ldots,s_C$. Each response can be classified into one of C response classes r_1, $r_2,\ldots,r_C$. This classification can be represented by the matrix $N(s_\alpha, r_\beta)$, whose entries denote the number of times a stimulus s_α elicits a response in class r_β.

Initially set all of the elements of the matrix $N(s_\alpha, r_\beta)$ to zero. For each stimulus class s_γ, calculate $d(S,s_\gamma)$, the average distance from S to the spike trains elicited in stimulus class s_γ as:

$$d\left(S,s_\gamma\right)=\left[\left\langle\left(D[q](S,S')\right)^z\right\rangle_{S'}\right]^{1/z},$$

where the operator $\langle\ \rangle_{S'}$ denotes the mean over all values of S', and S' is the set of all spike trains elicited by the stimulus class s_γ. Note that spike train self-comparisons are excluded from this calculation (i.e. $D[q](S,S)$). The spike train S is then classified into the response class r_β that minimizes the average distance $d(S, s_\gamma)$, and $N(s_\alpha, r_\beta)$ is incremented by 1. If the minimum is shared by k distances, each element is incremented by $1/k$. In the analyses presented here, z=-2.

The information entropy, H, is used to quantify the extent to which $N(s_\alpha, r_\beta)$ is not

random:

$$H=\frac{1}{N_{tot}}\sum_{\alpha,\beta}N\left(s_\alpha,r_\beta\right)\left[\log_2 N\left(s_\alpha,r_\beta\right)-\log_2\sum_a N\left(s_\alpha,r_\beta\right)-\log_2\sum_b N\left(s_\alpha,r_\beta\right)+\log_2 N_{tot}\right].$$

SM 2 Bias correction

In the methods section we describe the procedure for bias-correction of the information entropy estimates. The estimated bias was subtracted from the final estimates presented in this paper. For the interested reader, we include these estimates here. Using the D^{count} metric, the estimated bias was 0.03±0.007 (standard error) for the contrast configuration of experiment 1, the estimated bias was and 0.04±0.007 in the direction configuration of experiment 1. The bias estimates using D^{spike} were 0.1±0.007 and 0.12±0.008 in the contrast and direction configurations of experiment

1. The bias estimates using $D^{product}$ were 0.1±0.009 and 0.9±0.009 in the contrast and direction configurations of experiment 1.

We used the same technique to determine if the signal detection analysis was biased. However, after shuffling the metric estimates within the metric matrix (shown in figure 3), it was apparent that estimates were always very close to 0.50 (i.e. this was an unbiased estimator). Therefore no bias correction was done on the performance data. Specifically, using the D^{count} metric, the shuffled performance was 0.50±0.001 for the contrast and direction configurations of experiment 1. The shuffled performance values using D^{spike} were 0.5±0.002 in both the contrast and direction configurations of experiment 1. The shuffled performance values using $D^{product}$ were 0.5±0.001 in the contrast configuration and 0.5±0.002 in the direction configurations of experiment 1.

SM 3 Information entropy using all conditions

The analyses presented in this paper (Fig 4 – Fig 10) evaluated information entropy and the performance of a theoretical observer performing a task using two stimuli drawn from a set of seven total stimulus conditions. The stimuli were selected to maximize performance as long as it remains less than 82% (see methods section). This was done to test the theory that spike timing may be informative (and allow for improved performance) in situations in where a theoretical observer analysis using spike counts does not perform optimally. In this section, we present the information entropy of stimulus clustering comparing D^{count} and D^{spike} in which all 7 contrast and direction conditions of experiment 1 were used (Figure 1 SM). In experiment 2, we used first 7 conditions (motion directions separated by 30° intervals) to allow for comparison to the results from experiment 1. The theoretical observer analysis is omitted because its performance, at least in its present form (as defined in the appendix), is ill-defined for more than two conditions.

In the contrast configuration of experiment 1 (figure 1 SM - A), using the D^{count} metric, the mean entropy values in M1 and M2 were 0.57 and 0.69. When D^{spike} was used, the mean entropy values increased to 0.84 and 0.90 respectively (2-tail paired t-test, $p<0.001$ in both cases). This difference was significant in 37 cells (Wilcoxon rank-sum, $p<0.05$ using the bootstrap method). In the direction configuration of experiment 1 (figure 2 SM - B), using the D^{count} metric, the mean entropy values in M1 and M2 were 0.55 and 0.54. When D^{spike} was used, the mean

entropy values increased to 0.81 and 0.80 respectively (2-tail paired t-test: $p<0.001$ in both cases, significant increases in 44 cells using Wilcoxon rank-sum, $p<0.05$ with the bootstrap method). This trend was also seen in experiment 2 (figure 3 SM - C), in which using the D^{count} metric, the mean entropy values in M1 and M2 were 0.74 and 0.80. When D^{spike} was used, the mean entropy values increased to 0.95 and 1.00 respectively (2-tail paired t-test, $p<0.001$ in both cases, significant in 28 cells using Wilcoxon rank-sum, $p<0.05$ with the bootstrap method).

SM 4 $D^{product}$ performance in experiment 2

Figure 2 SM replots the data from figure 8. Values of information entropy (in A) and theoretical observer performance (in B) using the $D^{product}$ metric were omitted from this figure for brevity. This analysis is included here for the interested reader with comparisons to D^{count} and D^{spike}. In both M1 and M2, the information entropy was 0.62. This value was significantly larger than using that of D^{count} (2-tail paired t-test, $p<0.001$ in both animals), and was significantly smaller than D^{spike} in M2 (2-tail paired t-test, $p<0.001$) but was not significant in M1 (2-tail paired t-test, $p=0.13$). At the single unit level, 38 cells (out of 103) yielded higher information entropy using $D^{product}$ over D^{count}, compared to 14 cells in which information entropy was higher using D^{count} over $D^{product}$. 48 cells yielded higher entropy using D^{spike} over $D^{product}$ compared to 11 cells with higher information entropy using $D^{product}$ over D^{spike} (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method).

In the theoretical observer analysis (Figure 2 SM (B)), $P_{correct}=0.73$ and $P_{correct}=0.72$ in M1 and M2 respectively. These performance values were not statistically different than using D^{count} (2-tail paired t-test, $p=0.17$ and $p=0.9$), and were statistically lower than using D^{spike} (2-tail paired t-test, $p<0.001$ in both animals). The single unit analysis demonstrated improved theoretical observer performance in 24 cells using $D^{product}$ over D^{count} compared with 49 cells in which performance was better using D^{count} (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method). 60 cells produced higher levels of performance using D^{spike} over $D^{product}$ compared to 16 cells in which performance was better using $D^{product}$ (Wilcoxon rank-sum with $p<0.05$ using the bootstrap method).

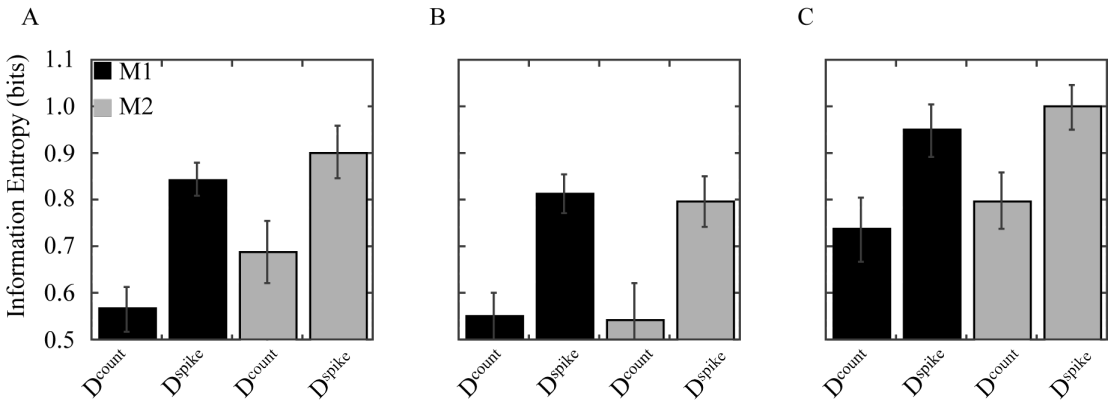


Figure 1 SM. Information entropy of stimulus clustering over all conditions in the contrast configuration of Experiment 1 (A), the direction configuration of Experiment 1 (B), and in Experiment 2 (C). The abscissa displays the metric and the ordinate the information entropy in bits. The colors represent data from different animals.

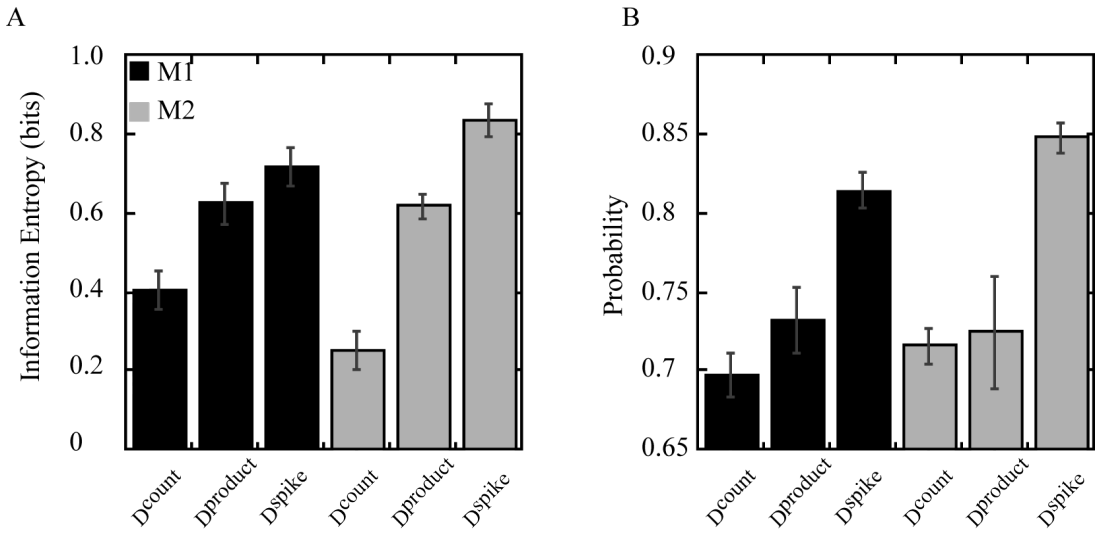


Figure 2 SM. Results of experiment 2. (A) Information entropy of stimulus clustering (ordinate) using the D^{count}, $D^{product}$ and D^{spike} metrics (abscissa). (B) Theoretical observer performance in a direction discrimination task (ordinate) using the D^{count}, $D^{product}$ and D^{spike} metrics (abscissa). The color represents data from different animals.

Chapter 3

Summary

The work included here examines cognitive influences on the processing of visual information both on the neural and the behavioral level. Recording single unit and LFP activity in area MT of the macaque monkey, we have gained deeper insights into the neural mechanisms of various aspects of visual selective attention that, so far, had been unexplored.

Specifically, we demonstrated that attention individually enhances representations of multiple moving objects. This enhancement is based on complex interactions between, the spatial separation of the attended objects, the size of individual neuron's RFs, and the specific role of a brain area in the task. Furthermore, we showed that attention modulates the input signals of MT neurons, and that these effects are reflected in certain frequency bands of LFP oscillations. Finally, we showed that spike timing is a source of information likely used by the brain for encoding different features of visual stimuli.

Complementing the electrophysiological studies, our behavioral experiments investigated consequences of dividing attention between stimuli represented in two different reference frames and the influence of feature-based attention on the processing of information from non-overlapping spatial locations. Collectively, these studies show that cognitive factors strongly modulate the processing of sensory information in primate visual cortex.

Bibliography

Albright, T. D., Desimone, R., and Gross, C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology,* 51(1): 16-31.

Allen, R., McGeorge, P., Pearson, D., and Milne, A. B. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive Psychology,* 18(3): 251-371.

Allman, J., Miezin, F., and McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience,* 8: 407-430.

Awh, E. and Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance,* 26(2): 834-846.

Barlow, H. B., Blakemore, C., and Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *Journal of Physiology,* 193(2): 327-342.

Bichot, N. P., Rossi, A. F., and Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science,* 308(5721): 529-534.

Blaser, E., Pylyshyn, Z. W., and Holcombe, A. O. (2000). Tracking an object through feature space. *Nature,* 408(6809): 196-199.

Born, R. T. and Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience,* 28: 157-189.

Brefczynski, J. A. and DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience,* 2(4): 370-374.

Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience,* 12(12): 4745-4765.

Buschman, T. J. and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science,* 315(5820): 1860-1862.

Carrasco, M., Ling, S., and Read, S. (2004). Attention alters appearance. *Nature Neuroscience,* 7(3): 308-313.

Cavanagh, P. and Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences,* 9(7): 349-354.

Cowey, A. and Stoerig, P. (1991). The neurobiology of blindsight. *Trends in Neurosciences,* 14(4): 140-145.

Culham, J. C., Brandt, S. A., Cavanagh, P., Kanwisher, N. G., Dale, A. M., and Tootell, R. B. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology,* 80(5): 2657-2670.

Cumming, B. G. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature,* 418(6898): 633-636.

Desimone, R. and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience,* 18: 193-222.

Desimone, R. and Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *Journal of Comparative Neurology,* 248(2): 164-189.

DeYoe, E. A. and Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences,* 11(5): 219-226.

Drew, T., McCollough, A. W., Horowitz, T. S., and Vogel, E. K. (2009). Attentional enhancement during multiple-object tracking. *Psychonomic Bulletin & Review,* 16(2): 411-417.

Dubner, R. and Zeki, S. M. (1971). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research,* 35(2): 528-532.

Egly, R., Driver, J., and Rafal, R. D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General,* 123(2): 161-177.

Eriksen, C. W. and St James, J. D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Perception & Psychophysics,* 40(4): 225-240.

Fecteau, J. H. and Munoz, D. P. (2005). Correlates of capture of attention and inhibition of return across stages of visual processing. *Journal of Cognitive Neuroscience,* 17(11): 1714-1727.

Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex,* 1(1): 1-47.

Gardner, J. L., Merriam, E. P., Movshon, J. A., and Heeger, D. J. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *Journal of Neuroscience,* 28(15): 3988-3999.

Gattass, R., Nascimento-Silva, S., Soares, J. G., Lima, B., Jansen, A. K., Diogo, A. C., et al. (2005). Cortical visual areas in monkeys: location, topography, connections, columns, plasticity and cortical dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 360(1456): 709-731.

Ghose, G. M. (2009). Attentional modulation of visual responses by flexible input gain. *Journal of Neurophysiology,* 101(4): 2089-2106.

Girard, P., Salin, P. A., and Bullier, J. (1992). Response selectivity of neurons in area MT of the macaque monkey during reversible inactivation of area V1. *Journal of Neurophysiology,* 67(6): 1437-1446.

Goodale, M. A. and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences,* 15(1): 20-25.

Heinze, H. J., Luck, S. J., Munte, T. F., Gos, A., Mangun, G. R., and Hillyard, S. A. (1994). Attention to adjacent and separate positions in space: an electrophysiological analysis. *Perception & Psychophysics,* 56(1): 42-52.

Henderson, J. M. and Macquistan, A. D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics,* 53(2): 221-230.

Hendry, S. H. and Yoshioka, T. (1994). A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus. *Science,* 264(5158): 575-577.

Hillyard, S. A. and Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences,* 95(3): 781-787.

Horton, J. C. and Hubel, D. H. (1981). Regular patchy distribution of cytochrome oxidase staining in primary visual cortex of macaque monkey. *Nature,* 292(5825): 762-764.

Howe, P. D., Horowitz, T. S., Morocz, I. A., Wolfe, J., and Livingstone, M. S. (2009). Using fMRI to distinguish components of the multiple object tracking task. *Journal of Vision,* 9(4): 10 11-11.

Kaas, J. H. and Lyon, D. C. (2001). Visual cortex organization in primates: theories of V3 and adjoining visual areas. *Progress in Brain Research,* 134: 285-295.

Komatsu, H. and Wurtz, R. H. (1988). Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *Journal of Neurophysiology,* 60(2): 580-603.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences,* 7(1): 12-18.

Lee, J. and Maunsell, J. H. (2009). A normalization model of attentional modulation of single unit responses. *PLoS One,* 4(2): e4651.

Leopold, D. A., Bondar, I. V., and Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature,* 442(7102): 572-575.

Ling, S. and Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research,* 46(8-9): 1210-1220.

Livingstone, M. and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science,* 240(4853): 740-749.

Martinez-Trujillo, J. and Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron,* 35(2): 365-370.

Maunsell, J. H., Nealey, T. A., and DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience,* 10(10): 3323-3334.

Maunsell, J. H. and Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences,* 29(6): 317-322.

Maunsell, J. H. and Van Essen, D. C. (1983a). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology,* 49(5): 1127-1147.

Maunsell, J. H. and Van Essen, D. C. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology,* 49(5): 1148-1167.

Maunsell, J. H. and van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience,* 3(12): 2563-2586.

Maunsell, J. H. and Van Essen, D. C. (1987). Topographic organization of the middle temporal visual area in the macaque monkey: representational biases and the relationship to callosal connections and myeloarchitectonic boundaries. *Journal of Comparative Neurology,* 266(4): 535-555.

McAdams, C. J. and Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience,* 19(1): 431-441.

McAlonan, K., Cavanaugh, J., and Wurtz, R. H. (2008). Guarding the gateway to cortex with attention in visual thalamus. *Nature,* 456(7220): 391-394.

McMains, S. A. and Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron,* 42(4): 677-686.

Moore, T. and Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature,* 421(6921): 370-373.

Moran, J. and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science,* 229(4715): 782-784.

Morawetz, C., Holz, P., Baudewig, J., Treue, S., and Dechent, P. (2007). Split of attentional resources in human visual cortex. *Visual Neuroscience,* 24(6): 817-826.

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology,* 70(3): 909-919.

Movshon, J. A. and Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience,* 16(23): 7733-7741.

Muller, M. M., Malinowski, P., Gruber, T., and Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature,* 424(6946): 309-312.

Muller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., and Brandt, S. A. (2003a). A physiological correlate of the "Zoom Lens" of visual attention. *Journal of Neuroscience,* 23(9): 3561-3565.

Nassi, J. J. and Callaway, E. M. (2006). Multiple circuits relaying primate parallel visual pathways to the middle temporal area. *Journal of Neuroscience,* 26(49): 12789-12798.

Newsome, W. T. and Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience,* 8(6): 2201-2211.

O'Craven, K. M., Downing, P. E., and Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature,* 401(6753): 584-587.

Pack, C. C., Hunter, J. N., and Born, R. T. (2005). Contrast dependence of suppressive influences in cortical area MT of alert macaque. *Journal of Neurophysiology,* 93(3): 1809-1815.

Pasternak, T. and Merigan, W. H. (1994). Motion perception following lesions of the superior temporal sulcus in the monkey. *Cerebral Cortex,* 4(3): 247-259.

Pasupathy, A. (2006). Neural basis of shape representation in the primate brain. *Prog Brain Research,* 154: 293-313.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology,* 32(1): 3-25.

Pylyshyn, Z. W. (2006). Some puzzling findings in multiple object tracking (MOT): II. Inhibition of moving nontargets. *Visual Cognition,* 14(2): 175-198.

Pylyshyn, Z. W., Haladjian, H. H., King, C. E., and Reilly, J. E. (2008). Selective nontarget inhibition in multiple object tracking (MOT). *Visual Cognition,* 16(8): 1011-1021.

Pylyshyn, Z. W. and Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision,* 3(3): 179-197.

Raiguel, S., Van Hulle, M. M., Xiao, D. K., Marcar, V. L., and Orban, G. A. (1995). Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. *European Journal of Neuroscience,* 7(10): 2064-2082.

Reynolds, J. H. and Heeger, D. J. (2009). The normalization model of attention. *Neuron,* 61(2): 168-185.

Reynolds, J. H., Pasternak, T., and Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron,* 26(3): 703-714.

Rizzolatti, G. and Matelli, M. (2003). Two different streams form the dorsal visual system: anatomy and functions. *Experimental Brain Research,* 153(2): 146-157.

Rizzolatti, G., Riggio, L., Dascola, I., and Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia,* 25(1A): 31-40.

Rodman, H. R., Gross, C. G., and Albright, T. D. (1990). Afferent basis of visual response properties in area MT of the macaque. II. Effects of superior colliculus removal. *Journal of Neuroscience,* 10(4): 1154-1164.

Roelfsema, P. R., Lamme, V. A., and Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature,* 395(6700): 376-381.

Saenz, M., Buracas, G. T., and Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience,* 5(7): 631-632.

Salzman, C. D., Murasugi, C. M., Britten, K. H., and Newsome, W. T. (1992). Microstimulation in visual area MT: effects on direction discrimination performance. *Journal of Neuroscience,* 12(6): 2331-2355.

Schein, S. J. and de Monasterio, F. M. (1987). Mapping of retinal and geniculate neurons onto striate cortex of macaque. *Journal of Neuroscience,* 7(4): 996-1009.

Schiller, P. H., Sandell, J. H., and Maunsell, J. H. (1986). Functions of the ON and OFF channels of the visual system. *Nature,* 322(6082): 824-825.

Shapley, R. and Perry, H. V. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences,* 9: 229-235.

Silver, M. A. and Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. *Trends in Cognitive Sciences,* 13(11): 488-495.

Sincich, L. C., Park, K. F., Wohlgemuth, M. J., and Horton, J. C. (2004). Bypassing V1: a direct geniculate input to area MT. *Nature Neuroscience,* 7(10): 1123-1128.

Snowden, R. J., Treue, S., Erickson, R. G., and Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience,* 11(9): 2768-2785.

Spitzer, H., Desimone, R., and Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science,* 240(4850): 338-340.

Tanaka, K., Hikosaka, K., Saito, H., Yukie, M., Fukada, Y., and Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *Journal of Neuroscience,* 6(1): 134-144.

Tovee, M.J. (2008*). An introduction to the visual system*. Cambridge University Press.

Treue, S. and Martinez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature,* 399(6736): 575-579.

Treue, S. and Maunsell, J. H. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature,* 382(6591): 539-541.

Ungerleider, L. G. and Mishkin, M. (1982). Two cortical visual systems. In Ingle, D. J., Goodale, M. A., and Mansfield, R. J. W., editors, *Analysis of Visual Behavior*, pages 549–586. MIT Press, Cambridge, MA.

Van Essen, D. C., Maunsell, J. H., and Bixby, J. L. (1981). The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *Journal of Comparative Neurology,* 199(3): 293-326.

Yantis, S. (1992). Multielement visual tracking: attention and perceptual organization. *Cognitive Psychology,* 24(3): 295-340.

Zhang, W. and Luck, S. J. (2009). Feature-based attention modulates feedforward visual processing. *Nature Neuroscience,* 12(1): 24-25.

Zohary, E., Shadlen, M. N., and Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature,* 370(6485): 140-143.

Curriculum Vitae

Robert Niebergall

Education

08/06–12/09	Visiting research trainee at McGill University, Montreal, Canada Advisor: Prof. Dr. Martinez–Trujillo
01/05–12/09	Ph.D.-student at the German Primate Center, Göttingen Thesis: Neuronal and perceptual effects of selective attention in the primate visual system. Advisor: Prof. Dr. S.Treue
11/03–12/04	Diploma in Biology, Georg-August Universität and German Primate Center, Göttingen Thesis: The spatial distribution of attention and its effect on contrast sensitivity. Advisor: Prof. Dr. S.Treue
09/02–07/03	Erasmus-student at the Universitat de Barcelona, Barcelona, Spain
10/00–08/02	Graduate studies in Biology, Georg-August Universität Göttingen
09/98–09/00	Undergraduate studies in Biology, Philipps-Universität Marburg
09/97–08/98	Voluntary social work at the Johanna-Kirchner Altenpflegezentrum, Frankfurt am Main
07/97	Abitur, Landschulheim Steinmühle, Marburg

Relevant Experience

09/05	Participant in the CNS Course on Computational Neuroscience, Max-Planck-Institut for Dynamics and Self-Organization, Göttingen.
09/05	Participant in the SensoPrim training program, German Primate Center, Göttingen.
07/05	Participant in the Seminario de Neurociencias: "Brain mechanisms of attention: biological and computational approaches", Soria, Spain.

11/04–01/5 Research Assistant
Cognitive Neuroscience Laboratory, German Primate Center, Göttingen

04/02–08/02 Research Assistant
Laboratory for Molecular Neurobiology, Max-Planck Institut für Experimentelle Medizin, Göttingen

04/01–04/02 Research Assistant (part-time)
Abteilung für Sensomotorik, Georg-August Universität, Göttingen

03/01 Internship
Abteilung für Metabolische Stoffwechselphysiologie, Georg-August Universität, Göttingen

Scholarships

06/07–05/08 Scholarship by the German Academic Exchange Service (DAAD)

Journal Publications

Tzvetanov, T., Womelsdorf, T., Niebergall. R., Treue, S., (2006). Feature-based attention influences contextual interactions during motion repulsion. *Vision Research*, 46(21): 3651–3658.

Hardeland, R., Poeggeler, B., Niebergall, R., Zelosko, V., (2003). Oxidation of melatonin by carbonate radicals and chemiluminescence emitted during pyrrole ring cleavage. *Journal of Pineal Research*, 34(1): 17-25.

Published abstracts

Niebergall, R., Khayat, P.S., Treue, S., Martinez-Trujillo, J.C. (2009). Splitting and zooming the focus of attention in primate visual cortex during multiple-object tracking. Poster presented at the 39th annual meeting, Neuroscience, Chicago, USA.

Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C. (2009). Feature-Based Attention Modulates the Strength of Direction-Selective Inputs into MT Neurons. Poster presented at the 39th annual meeting, Neuroscience, Chicago, USA.

Niebergall, R., Khayat, P.S., Treue, S., Martinez-Trujillo, J.C. (2009). Splitting the spotlight of attention during multiple-object tracking. Poster presented at the 32nd Göttingen Neurobiology Conference, Göttingen.

Niebergall, R., Khayat, P.S., Treue, S., Martinez-Trujillo, J.C. (2009). Multiple attentional spotlights modulate the activity of single neurons in visual cortex during multiple object tracking. Talk presented at the winter meeting of the Canadian Physiological Society, Beaupré, Québec, Canada.

Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C. (2009). Modulation of local field potentials in different frequency bands reflects bottom-up and top-down effects of attention in middle temporal visual area. Talk presented at the winter meeting of the Canadian Physiological Society, Beaupré, Québec, Canada.

Sachs A.J., Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C. (2009). Temporal coding of visual features by neurons in middle temporal area. Talk presented at the winter meeting of the Canadian Physiological Society, Beaupré, Québec, Canada.

Niebergall, R., Khayat, P.S., Treue, S., Martinez-Trujillo, J.C. (2008). Neural correlates of multiple spotlights of attention in area MT during multiple-object tracking. Talk presented at the 38th annual meeting, Neuroscience, Washington, USA.

Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C. (2008). Attentional modulation of local field potentials response tuning to direction and contrast in macaque's area MT. Talk presented at the 38th annual meeting, Neuroscience, Washington, USA.

Khayat, P.S., Niebergall, R., Martinez-Trujillo, J.C. (2007). Spatial attentional modulation in area MT depends on visual inputs. Talk presented at the 37th annual meeting, Neuroscience, San Diego, USA.

Niebergall, R., Martinez-Trujillo, J.C. (2007). Reference frames for covert spatial attention during smooth pursuit tracking of visual targets. Poster presented at the 7th annual meeting, Vision Sciences Society, Sarasota, Florida, USA.

Niebergall, R., Martinez-Trujillo, J.C. (2007). Reference frames for covert spatial attention during a visuomotor task. Poster presented at the 31st Göttingen Neurobiology Conference, Göttingen.

Folta, K., Niebergall, R. Fischbach, A., Grube, D., Treue, S. (2007). Stimulus modality differentiates human and non-human timing and memory of rythmic light and tone signals. Poster presented at the 31st Göttingen Neurobiology Conference, Göttingen.

Niebergall, R., Martinez-Trujillo, J.C. (2007). Covert attention tracks visual targets in eye-centered coordinates. Talk presented at the winter meeting of the Canadian Physiological Society, Beaupré, Québec, Canada.

Niebergall, R., Tzvetanov, T., Treue, S. (2005). The spatial profile of attentional modulation of contrast detection thresholds. Poster presented at the 35th annual meeting, Neuroscience, Washington, USA.

Niebergall, R., Tzvetanov, T., Treue, S. (2005). The effect of allocating spatial attention on contrast thresholds across the visual field. Poster presented at the 1st Meeting of the 4 German BCCNs and Inauguration of the BCCN Freiburg, Freiburg.

Niebergall, R., Tzvetanov, T., Treue, S. (2005). The effect of allocating spatial attention on contrast thresholds across the visual field. Poster presented at the European Conference on Visual Perception, A Coruña, Spanien.

Niebergall, R., Womelsdorf, T., Tzvetanov, T., Treue, S. (2005). The spatial distribution of attentional facilitation/inhibition of contrast sensitivity depends on task context. Poster presented at the 30th Göttingen Neurobiology Conference, Göttingen.

Göttingen, 8th of December, 2009.