2015

Specialized Signals for Spatial Attention in the Ventral and Dorsal Visual Streams

Pablo Polosecki

SPECIALIZED SIGNALS FOR SPATIAL ATTENTION IN THE VENTRAL AND DORSAL VISUAL STREAMS

A thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Pablo Polosecki

June 2015

SPECIALIZED SIGNALS FOR SPATIAL ATTENTION IN THE VENTRAL AND DORSAL VISUAL STREAMS

Pablo Polosecki, Ph.D.

The Rockefeller University 2015

Neuroscientists have traditionally conceived the visual system as having a ventral stream of vision for perception and a dorsal one associated with vision for action. However functional differences between them have become relatively blurred in recent years, not the least by the systematic parallel mapping of functions allowed by functional magnetic resonance imaging (fMRI).

Here, using fMRI to simultaneously monitor several brain regions, we first studied a hallmark ventral stream computation: the processing of faces. We did so by probing responses to motion, an attribute whose processing is typically associated with the dorsal stream. In humans, it is known that face-selective regions in the superior temporal sulcus (STS) show enhanced responses to facial motion that are absent in the rest of the face-processing system. In macaques, face areas also exist, but their functional specializations for facial motion are unknown. We showed static and moving face and non-face objects to macaques and humans in an fMRI experiment in order to isolate potential functional specializations in the ventral stream face-processing system and to motivate putative homologies across species. Our results revealed all macaque face areas

showed enhanced responses to moving faces. There was a difference between more dorsal face areas in the fundus of the STS, which are embedded in motion responsive cortex and ventral ones, where enhanced responses to motion interacted with object category and could not be explained by their proximity to motion responsive cortex. In humans watching the same stimuli, only the STS face area showed an enhancement for motion. These results suggest specializations for motion exist in the macaque face-processing network but they do not lend themselves to a direct equalization between human and macaque face areas.

We then proceeded to compare ventral and dorsal stream functions in terms of their code for spatial attention, whose control was typically associated with the dorsal stream and prefrontal areas. We took advantage of recent fMRI studies that provide a systematic map of cortical areas modulated by spatial attention and suggest PITd, a ventral stream area in the temporal lobe, can support endogenous attention control. Covert attention and stimulus selection by saccades are represented in the same maps of visual space in attention control areas. Difficulties interpreting this multiplicity of functions led to the proposal that they encode priority maps, where multiple sources are summed to form a single priority signal, agnostic as to its eventual use by downstream areas. Using a paradigm that dissociates covert attention and response selection, we test this hypothesis with fMRI-guided electrophysiology in two cortical areas: parietal area LIP, where the priority map was first proposed to apply, and temporal area PITd.

Our results indicate LIP sums disparate signals, but as a consequence independent channels of spatial information exist for attention and response planning. PITd represents relevant locations and, rather than summing signals, contains a single map for covert attention. Our findings have the potential to resolve a longstanding controversy about the nature of spatial signals in LIP and establish PITd as a robust map for covert attention in the ventral stream.

Together, our results suggest that while the distribution of labor between ventral stream and dorsal stream areas is less linear than what a what a rough depiction of them can suggest, it is illuminated by their proposed function as supporting vision for perception and vision for action respectively.

To Mechi

Acknowledgements

I am indebted to Winrich for his committed mentoring. He would always be available for his students, and glad to discuss ideas, especially when the discussion had a conceptual motivation. Beyond scientific knowledge and technique, he taught me the importance of remaining faithful to one's original convictions and goals in the face of difficulties, and insisted that I should strive to generate something of which I could be proud. I have seen him create a lab from scratch and run it with total dedication, yet always open to change plans when a new idea seemed promising. I found in that mixture of flexibility and perseveration the true meaning of a doctorate.

My Thesis Committee members provided invaluable guidance, especially in those moments of inevitable frustration and despair. Cori Bargmann and Charles Gilbert found time to talk and give advice when I really needed it. Their wisdom, from purely scientific questions to more general considerations about scientific life, was vital.

I was fortunate enough to work side by side with Sara Steenrod. Sara's sense of humor, ability to understand people and animals alike and her excitement for the secret music of neurons was an invaluable company during the harsh times of making experiments work. Some of those conversations in the lab should be turned into a book, although one for a very specific audience.

I am also grateful to the great people that make Winrich's lab. It is hard to reach such a delicate balance of people willing to discuss ideas, provide pieces of computer code, take care of each other's animals, and simply give a helping hand expecting nothing in return. Ours is a kind of research that simply cannot happen without generous colleagues, and you guys certainly made it happen. I am proud to be a member of this group.

I would also like to thank the members of the Tsao lab, who first taught me how to do experiments. They are the only thing I miss from Pasadena!

I must mention the impeccable job done by veterinaries and husbandry staff. Their professionalism and interest in the well being of the animals (and scientists!) is behind all these experiments.

Beyond science, family and friends in Buenos Aires and New York City played an essential role during this time, reminding me that there is a world out there of people who don't have a PhD, and yet lead happy lives, love each other and have interesting things to say. They were my anchor to a reality of which one can surprisingly lose sight in science for moments.

Finally, there is the one person who made it worthwhile to jump down the rabbit hole, my wife Mercedes. She not only gave every daily little experience a new meaning, but was never afraid to throw everything else away, so we could be together while I finished my thesis. My PhD truly started when she appeared, and now it is time to find out what our next adventure will be.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

The core function of the brain is to guide behavior based on internal goals and judgments about the state of the world (Braitenberg, 2007). In the visual system, these different aspects of brain function (the guiding of behavior and construction of high-level judgments) correspond to the distinction between fine-grained representations of the world (i.e., perception) and visually guided behavior. For decades now, neuroscientists have widely considered this duality of function to be mirrored by an anatomical division: that of a ventral and a dorsal visual stream. In this thesis we examine the processing of faces, a paradigmatic ventral stream computation, and discover specializations within the face-processing network using moving stimuli, an attribute traditionally associated with the dorsal visual stream. Then, we study specializations in representative areas from both streams for a function that has been ascribed to the dorsal stream (and other structures beyond both): the control of spatial attention.

Two streams of visual processing

Deeply rooted in our understanding of the visual system, the distinction between a ventral, occipitotemporal, and a dorsal, occipitoparietal, cortical stream was originally proposed by Ungerleider and Mishkin (1982), based on neuroanatomical, electrophysiological and lesion studies. Initially understood as a dichotomy between the representation of features and identity of visual stimuli ("what") and their location ("where"), it was later conceptually reframed as supporting vision for perception and vision for action (Goodale and Milner, 1992). Given the complex and diverse connectivity of the many areas in these pathways, it is difficult to describe the general function of an entire pathway. Here we simply provide a high-level intuition about the general functions of both pathways.

Ventral stream processing starts in area V1 and extends anteriorly through areas V2 and V4 to the temporal lobe, including areas of inferotemporal (IT) cortex. As information processing flows towards anterior IT, receptive field sizes increase and responses become selective to more complex stimuli, while at the same time acquiring tolerance to variations in luminance, point of view, etc. What is common to these structures is their involvement in forming specific representations or associations comprising stable aspects of visual information. The key aspect of ventral pathway representations is not that they are tied to

particular physical objects, but that they capture a stable configuration of visual information (e.g., texture, scenes) (Kravitz et al., 2013).

The dorsal pathway is an occipitoparietal network that lies between the early visual cortex and specialized cortical structures involved in visually guided action, spatial attention, somatosensation, spatial aspects of audition, navigation, and spatial working memory. The occipitoparietal circuit consists of a set of projections from early visual cortical areas to posterior regions of the parietal cortex. It goes from portions of V1 to area V6 (dorsomedial area), which receives inputs from areas V2, V3, V3a. From V6 there is a medial set of projections to areas V6a, MIP and VIP while another projection crosses laterally to areas MT, MST and LIP. This circuit is the common origin to three distinct subpathways that branch from there: a parietofrontal pathway from areas LIP, VIP, MT and MST; a parieto-premotor pathway from areas 6A and MIP and a parieto-medial temporal pathway from the caudal inferoparietal lobule, which is the most complex of the three. A detailed discussion of these can be found in a recent review by Kravitz and colleagues (2011). Neurons within the dorsal stream respond selectively to spatial aspects of stimuli, such as the direction and speed of stimulus motion. Such cells also respond when the animal visually tracks a moving target (Ungerleider and Pessoa, 2008). Importantly, cells respond to the movements of effectors such as the eyes and hands. More generally, the occipitoparietal network must create a map of the relative positions between visual stimuli and effectors to guide action effectively. The need to represent these sorts of

relationships naturally leads to the formation of coordinate systems and general reference frames. In this way, the dorsal pathway specializes in capturing arbitrary and dynamic spatiotemporal relationships between multiple items (Kravitz et al., 2013).

It should be noted that deviation from even these broad characterizations have been found in the dorsal stream. Shape selectivity in the dorsal stream has been reported in macaques (e.g., Sereno and Maunsell, 1998) and in humans (e.g., Konen and Kastner, 2008). The latter findings are particularly remarkable, as they suggest not only that shape selectivity can be found there but that a true hierarchy of objects representation exits there with increasing tolerance to stimulus transformations in sequential processing stages reminiscent of those found in the ventral stream. Conversely, some areas in the ventral stream also show attention signals that are reminiscent of those in the dorsal stream. We discuss this in detail in subsequent sections.

Ventral stream computations: a study on facial motion

A paradigmatic example of ventral stream computation is the processing of face stimuli, for which dedicated cortical regions exist in humans (Kanwisher et al., 1997) and macaques (Tsao et al., 2003; 2006), where neural responses in successive stages of processing are increasingly invariant to changing aspects of

the same face and become increasingly selective to facial identity (Freiwald and Tsao, 2010).

Before moving to the direct comparison between signals from both streams, we will present in Chapter 2 a study showing how the analysis of motion responses, an attribute typically associated with dorsal stream specializations, can clarify our understanding of ventral stream computations in the context of face processing and motivate homologies between human and non-human brain regions.

Faces provide a rich source of social information. Some information, such as individual identity, is transmitted by the structure of the face. Other information, such as its mood, involves dynamic transformations (Darwin, 1872). Because of this, face recognition requires motion to be factored out for identification while simultaneously extracted to perceive changes in expression, head orientation, or gaze. The mechanisms for performing these very different computations have been suggested to reside in different parts of the human brain (Bruce and Young, 1986; Haxby et al., 2000; O'Toole et al., 2002). In particular, in human cortex, it has been suggested that the occipital face area (OFA) and the fusiform face area (FFA) represent invariant properties of faces (Kanwisher et al., 1997; McCarthy et al., 1997; Yovel and Kanwisher, 2004), whereas the superior temporal sulcus face area (STS-FA) is sensitive to dynamic face properties (Allison et al., 2000; Gobbini et al., 2011). Recently, Pitcher and colleagues (2011) found a clear functional dissociation, with the STS-FA selective for dynamic information and OFA and FFA insensitive to facial motion. As mentioned

above, in macaque monkeys, a network of face-selective areas has also been identified (Pinsk et al., 2009; Rajimehr et al., 2009; Tsao et al., 2003; 2008), but specializations for facial motion have not been investigated yet. To better understand how facial motion is processed across species, we probed the face-processing networks of both macaques and humans to address two questions: is the processing of dynamic information functionally separated within face-processing networks? If so, how does this separation inform putative homologies of face areas across the two primate species? We present the results of these experiments in chapter 2.

Spatial Attention: perception and eye movements

The bulk of the thesis is devoted to the study of spatial attention. While a unified definition of attention is elusive, the term refers to the selection of a subset of the available visual information for enhanced processing (Ward, 2008), which is thought to be a necessary condition for making the problem of vision computationally tractable (Tsotsos, 2011). While often this is made by orienting sensory organs to relevant stimuli, as in the case of saccadic eye movements, it is possible to attend to a location or stimulus covertly as first studied by Helmholtz (1867). The relevant entities or dimensions selected by attention can be spatial locations, visual features or entire objects. Here, we will focus on spatial attention: the selection of a location either covertly for perception or to

overly orient toward it. Behavioral effects of spatial attention include reduced reaction times (Posner, 1980), facilitation of memory (Sperling, 1960), enhancement of visual discrimination (Lee et al., 1997) and gating the detection of stimuli that would not otherwise reach awareness (Rensink, 2002). These effects have crucial practical consequences for the study of attention. Just like saccades are an observable consequence of motor planning, measuring the existence of a perceptual advantage behaviorally for stimuli being covertly attended can operationalize covert attention.

Covert attention is closely linked to the planning of eye movements. Both share part of their associated neural substrates including the superior colliculus (SC) (Knudsen, 2012), frontal eye fields (FEF) (Armstrong et al., 2012) and area LIP (Goldberg et al., 2012). Behaviorally, while covert attention can be regularly dissociated from eye movements (Posner, 1980), evidence indicates that right before an eye movement there is mandatory allocation of attention at the location of an impending saccade (Deubel and Schneider, 1996; Hoffman and Subramaniam, 1995; Kowler et al., 1995). Evidence of this relationship led to the conception of a premotor theory of attention, which proposed that “the program for orienting attention either overtly or covertly is the same, but in the latter case the eyes are blocked at a certain peripheral stage” (Rizzolatti et al., 1987). However causal interventions in the FEF found a temporal dissociation between attention selection and saccade preparation (Juan et al., 2008; 2004), confirmed by behavioral measurements of attention (Smith and Schenk, 2007). Such

findings indicate saccade preparation is not a necessary component of covert attention, discrediting the equivalence between attention and motor planning (Smith and Schenk, 2012).

It should be noted, however, that many studies trying to dissociate effects of attention and saccades have used an anti-saccade paradigm in an effort to isolate a neural attention enhancement that is not explained by saccade planning (e.g., Gottlieb and Goldberg, 1999; Zhang and Barash, 2000) or vice versa (Steinmetz and Moore, 2014), where monkeys are given a previous instruction to saccade away from a cued location. However, cued locations and saccade direction in such paradigms are not independent but rather anticorrelated, and it is conceivable there are effects representing the queuing of an attention shift from cued location to saccade location rather than different saccadic and attention signals or conversely, that activity at the cued location is used to signal the initiation of an anomalous motor plan. An interesting variation, designed to study cognitive conflict due to stimulus-response incompatibly in FEF, was used by Sato and Schall (2003), where monkeys judged the orientation of a color singleton stimulus by making a saccade to it or an antisaccade. They reported two types of neurons: one selected the singleton stimulus initially and later was selective for saccade direction while the second one was only selective for saccade direction. That dissociation between stimulus selection and motor planning is sequential rather than simultaneous. We will return to this point when discussing our experimental design in Chapter 3.

Control of spatial attention: Lessons from area LIP

We have a relatively rich understating of the effects of attention in areas that encode stimulus features, a large fraction of which have been captured by unifying conceptual models (Lee and Maunsell, 2009; Reynolds and Heeger, 2009). It has been harder to conceptualize how different control areas (LIP, FEF, SC) select stimuli for enhanced processing. The issue is made more difficult considering that neural selection can be used for either covert attention or to guide eye movements.

A case in point is parietal area LIP. LIP is connected to dorsal and ventral and stream areas (Neal et al., 1988; Seltzer and Pandya, 1980) and also the oculomotor system (Andersen et al., 1990; Asanuma et al., 1985; Schall et al., 1995). Accordingly, visual and saccadic responses exist. A hallmark property of this area is revealed by its involvement in memory-guided saccades (MGSs): a type of task where the flashing of a visual stimulus must be followed by a saccade to the stimulus location after a variable period of time during which there is no visual stimulation (memory period). Cells in area LIP respond in a spatially selective manner not only when the visual stimulus is presented inside their receptive field (RF) and when the saccade is made, but also during the memory period (Gnadt and Andersen, 1988). The presence of tuned activity throughout the memory period is of interest because it does not reflect a response to the immediate visual stimulation or the execution of a motor program but the

maintenance in memory of a link between both. Nevertheless, the visually stimulated location, the one held in memory and the one to which the saccade is executed are the same, making it difficult to establish a univocal interpretation of the spatial signal being encoded. Since then, a number of studies suggested LIP activity to reflect saccadic intention (Barash et al., 1991a; 1991b; Colby et al., 1996; Snyder et al., 1997) or visual attention (Balan and Gottlieb, 2006; 2009; Balan et al., 2008; Bisley and Goldberg, 2003; Colby et al., 1996; Gottlieb et al., 1998; Oristaglio et al., 2006; Peck et al., 2009). This led to a prolonged disagreement about the functional role of LIP (Andersen et al., 1997; Andersen and Buneo, 2002; Bisley and Goldberg, 2010; Colby and Goldberg, 1999).

We discuss here two studies of particular relevance to the way signals in LIP are currently understood. In the first one, Bisley and Goldberg (2003) used a clever experimental design in which monkeys had to perform a perceptual decision embedded within an MGS task. In this version of the task monkeys had to decide if they were to execute or cancel a planned saccade to the target location by performing a discrimination of a second flashed stimuli. By presenting the stimulus to be discriminated at different times and locations, the authors could measure the location of attention in time and space by measuring changes in discrimination performance. In this way, behavioral measures of attention location could be compared to neural activity. They found that attention is indeed directed at the location of the MGS during the memory period. In addition they found that flashing a distractor elsewhere briefly draws attention away from the

future saccade location. The crucial finding is that neural activity in LIP predicts the time course of the capture of attention by the distractor. As long as the average activity elicited by the distractor (in neurons whose RFs encompass the distractor location) exceeds the average activity of the cells whose RFs are at the location of the target, attention is drawn to the distractor location. When the average activity elicited by the distractor falls below that of the cells that signal target location, attentional advantage is restored to the target location. This suggests allocation of attention is determined in the map created by LIP through a winner-take-all mechanism, such that the location of maximum activity signals that of attention.

The second relevant study is a reanalysis of responses during a free viewing visual search task (Ipata et al., 2009). There, the authors found that the activity of LIP cells can be modeled as a sum of several disparate factors: a visual component, a saccadic component and a “cognitive” component that indicates the presence of the searched target in the RF of cells. Interestingly, the cognitive component seems to interact with the saccadic one, which could be a crucial violation of the proposed encoding scheme as a simple sum. In addition, attention is not controlled by the behavioral paradigm, and the proposed encoding scheme awaits confirmation from a parameterized experimental design. This motivates our experiment in Chapter 3.

The preceding experiments, together with the difficulties trying to separate attention and saccadic signals, have led to the proposal that LIP constructs a

priority map (Bisley and Goldberg, 2010). The idea of a priority map draws from the older idea of salience maps (Itti and Koch, 2001), which suggested that a useful map for stimulus-driven attention can be achieved by combining signals from different visual feature maps. The term “priority” is used to indicate that, in addition to stimulus-driven factors, top-down signals reflecting internal goals and expectations contribute to the resulting map. The idea is that disparate bottom-up and top-down signals are summed to create a single agnostic signal whose peak can be used by visual areas to guide attention for perception and by the oculomotor system to guide saccades when appropriate. How and when this signal is used for attention or for saccades is left unspecified.

It is tempting to discard these ideas as merely sematic arguments. However they make testable predictions in a context where a monkey needs to plan a saccade based on evidence provided by a covertly relevant stimulus in an independent location. In this situation, a priority map is expected not to be able to distinguish the covertly relevant location from the one to which the saccade is being made. This is because at the core the priority map is the idea that the signals have no intrinsic meaning but rather it is elsewhere (in feature coding areas, or in the oculomotor system) that it is decided how to use the priority signal. Because of this, independent and simultaneous decoding of saccade direction and covertly attended locations should be impossible. This idea can be contrasted to a “pure intention” map, where signals indicate the plan to make an eye movement or a “pure attention” map, which indicates the currently attended

location. In a “pure intention” map, for instance, signals regarding eye movements could always be decoded, even if the saccade is made to a location that is already attended covertly. In this scenario, highlighting of covertly attended locations would occur if there is an accompanying motor plan but it would disappear as soon as there is a plan to make a saccade elsewhere. Conversely, a “pure attention” map would robustly reflect covertly attended locations. Eye-movement related signals are to be expected, but only when saccades are not made to a location already attended covertly, requiring an accompanying shift of attention.

Our experiment in Chapter 3 examines the priority map idea explicitly: we test if disparate signals are summed and whether there is a single quantity represented or information about saccades and attention is present and explicit. We contrast this with the predictions of a “pure attention” and a “pure intention” scenario.

Attention signals in V4: similarities and differences with a priority map

It is interesting to contrast these observations to those made in ventral area V4. The effects of attention in V4 have been studied for decades (Moran and Desimone, 1985). Cells in V4 display differential responses when a relevant object falls inside their RF as well as to an impending saccade (Bichot et al., 2005), to the point that it has been proposed that they form a ‘salience map’

(Mazer and Gallant, 2003). Yet, cells don't simply compute salience or behavioral relevance of locations but are selective to visual patterns, early in their responses. As such, the modulations of spatial attention are visually gated (Ekstrom et al., 2008; Moore and Armstrong, 2003) or modest baseline modulations (Luck et al., 1997) and depend on stimulus features (Bichot et al., 2005). Moreover, causal interventions in V4 don't affect spatial attention (Dagnino et al., 2015) and their effects are not attention-dependent (Merigan, 2000). Because of this, V4 doesn't fulfill key criteria (Fecteau and Munoz, 2006) for priority maps or areas that have a causal effect in spatial attention control: even though its position in the ventral streams potentially allows V4 to integrate information from multiple feature coding areas, its responses are not in general feature-independent and causal manipulations do not manipulate attention location.

A recent report found a selectivity to task relevant visual patterns in V4 that are reminiscent of LIP responses (Ipata et al., 2012) in a visual search task. This selectivity for task relevant targets is strictly task dependent and occurs later and independently of the classical tuning of V4 cells. Importantly, this effect is observed in tasks requiring feature attention (such as visual search) and not ones requiring spatial attention (such as a delayed saccade task). Because of that it has been proposed this kind of effect can support feature attention specifically but not spatial attention.

Area PITd: a ventral pathway area for the control of spatial visual attention?

Until recently, electrophysiology studies have investigated spatial attention signals in a number of brain areas, but a systematic mapping of their effects in the macaque brain was lacking. Recently, Stemmann and Freiwald (2015, manuscript in preparation) studied the effects of covert spatial attention in the macaque brain using fMRI. They trained monkeys to perform an attention-demanding covert motion discrimination task, related to the one we use in Chapter 3. Monkeys were given a cue to covertly attend one out of two moving-dot surfaces (MDSs), positioned in the left and right hemifields along the horizontal meridian. Animals reported the direction (eight possibilities) of a prolonged motion event in the cued surface, with an eye movement to one out of eight possible saccade targets. The experiment relies on the contralateral representation bias in visual areas. Based on it, a contrast of activity on attend-left vs. attend-right conditions was used to create a statistical map of attention effects. The experiment revealed effects in familiar areas but also others, like PITd that had not been mapped.

Located in inferotemporal cortex, area PITd belongs to the ventral visual stream, surrounded by areas selective stimulus visual features and objects. It is a relatively little studied area (Conway et al., 2007; Hikosaka, 1997; 1998; Kolster et al., 2014). Electrophysiological recordings in PITd using the same task show

dramatic effects of covert attention in PITd (Stemmann and Freiwald, 2015, manuscript in preparation). In addition, PITd cells are not selective to motion direction and the strong attention effects are also observed in a variation of the task using color instead of motion, to which PITd cell are also not selective. In addition, mapping of shape selectivity show little tuning to object categories or shape (but see Kolster et al., 2014). Moreover, causal interventions using electrical microstimulation during the task revealed effects consistent with a manipulation of covert attention location.

Therefore PITd satisfies key criteria (Fecteau and Munoz, 2006) for areas that control visual attention: attention strongly modulates responses, there is little to no tuning to visual features, causal manipulations influence spatial attention behaviorally and the area is positioned so as to reciprocally connect to neighboring feature-coding areas. These findings are in stark contrast with the responses observed in neighboring area V4 discussed above. Taken together, the findings suggest the intriguing possibility that PITd is in control of at least certain spatial attention signals. In Chapter 3 we test the hypothesis that PITd implements a priority map as suggested for other attention control areas.

Dimensionality of neural representations

Areas that control attention are in a characteristic middle ground between sensory areas where important basic principles of organization are known (tuning

curves, ordered cortical maps) and “higher” cognitive areas, for which we lack a conceptual framework that would enable straightforward interrogation of the functional principles (Kiani et al., 2015). In part because of this and partially because of the flexibility required from them, activity in higher cognitive areas is known to display significant heterogeneity (Asaad et al., 1998; Mansouri et al., 2006). It has been proposed that in such areas, information is distributed across the population and that encoding benefits from high-dimensional representations caused by non-linear mixed selectivity (Rigotti et al., 2013). High-dimensionality manifests itself in cell responses as non-linear interactions between task variables and enables learning arbitrary combinations between them through simple linear decoders such as support vector machines (in fact the equivalence between non-linear decoding methods and linear ones in a higher-dimensional feature space is part of the standard bag of tricks in the field on machine learning).

In a number of ways this picture resembles the case of areas like LIP where heterogeneity of activity has been emphasized (Meister et al., 2013; Premereur et al., 2011) and responses to many disparate signals, spatial and not, are regularly reported (Gottlieb and Snyder, 2010). The idea of a priority signal constructed from the sum of all these sources tries to bring together those diverse findings in a manner consistent with modeling efforts suggesting LIP population dynamics to occur along a single dimension (Ganguli et al., 2008). In

the our discussion chapter we will return to these points in the light of the results presented in chapter 3.

What follows then is an investigation of visual processing along two complementary lines of research. First, we present a study on a canonical ventral stream computation: the processing of facial information. We show specialization within the face processing hierarchy to facial motion, which provide some clues as to potential homologies across species. Second, we present a study testing the priority map idea for the control of spatial attention. Using a paradigm that dissociates covert attention and response selection, we test this hypothesis with fMRI-guided electrophysiology in two areas: parietal area LIP, where the priority map proposal originated, and temporal area PITd, which recent evidence suggests can control covert attention. Our results indicate LIP sums disparate signals, but as a consequence independent channels of spatial information exist for attention and response planning. PITd represents relevant locations and, rather than summing signals, contains a single map for covert attention. We believe our findings resolve a longstanding controversy about the nature of spatial signals in LIP and establish PITd as a robust map for covert attention in the ventral stream.

Chapter 2

Specializations within ventral face processing hierarchy: the case of facial motion

In this chapter we present results of functional magnetic resonance imaging (fMRI) experiments in humans and macaques designed to investigate specializations of facial motion within the face-processing system. As mentioned in Chapter 1, in macaque monkeys, a network of face-selective areas has been identified, but specializations for facial motion have not been investigated. Since we know in humans specializations for facial motion reveal functional differences across face-selective areas, we probed the face-processing networks of both macaques and humans to address two questions: is the processing of dynamic information functionally separated within face-processing networks? If so, how does this separation inform putative homologies of face areas across the two primate species?

We scanned 3 macaque monkeys and 6 human subjects. In a first experiment, we used a standard set of static stimuli to localize face patches by contrasting responses to static faces with responses to static non-face objects (Experiment 1, Figure 2.1*a*, see Materials and Methods). In experiment 2, we probed the face-

processing system with visual stimuli consisting of blocks of face and non-face objects, either moving or static. To determine the position of face areas relative to motion-selective cortex, we derived maps of general motion sensitivity from Experiment 2 by contrasting responses to moving and static objects (Figure 2.1*b*, see Materials and Methods). In macaques, motion selectivity extended throughout the fundus of superior temporal sulcus, embedding the two face patches in the fundus of the STS, MF, and AF, in motion-responsive cortex (Figure 2.1*b*), whereas middle face patch on the STS lip (ML), anterior face patch on the STS lip (AL), and anterior face patch on the ventral surface of inferior temporal (AM) were not. A contrast of moving faces versus moving objects reproduced the known face patches (Figure 2.1*c*).

Figure 2.1

Face and motion selectivity in the macaque temporal lobe.

(a) (left) Face-selective regions in one representative hemisphere (monkey M1, right), on a flattened cortical surface. The color bar indicates the negative common logarithm of the probability of error. PL, posterior face patch; MF, middle face patch in the STS fundus; ML, middle face patch on the STS lip; AF, anterior face patch in the STS fundus; AL, anterior face patch on the STS lip; AM, anterior face patch on the ventral surface of IT. Sulci: sts, superior temporal; sf, sylvian fissure. (right) Time courses of fMRI signal for two representative regions (MF and ML). Colored epochs distinguish stimulation blocks. Block-types are indicated with symbols below the time axis for clarity (Green arrows: moving stimuli, stop sign: static stimuli, human or monkey faces: face stimuli, leaf or show: non-face object stimuli, see Materials and Methods for details)

(b) Map of the strength of motion responses. Face selective regions are represented by black outlines. The color bar indicates the magnitude of the response to motion (difference between the response to moving and static non-face objects) in units of percent signal change.

(c) Face selectivity map, similar to (a), but comparing moving faces to moving objects in experiment 2. Black outlines as in (b). AM falls partially outside the functional volume.

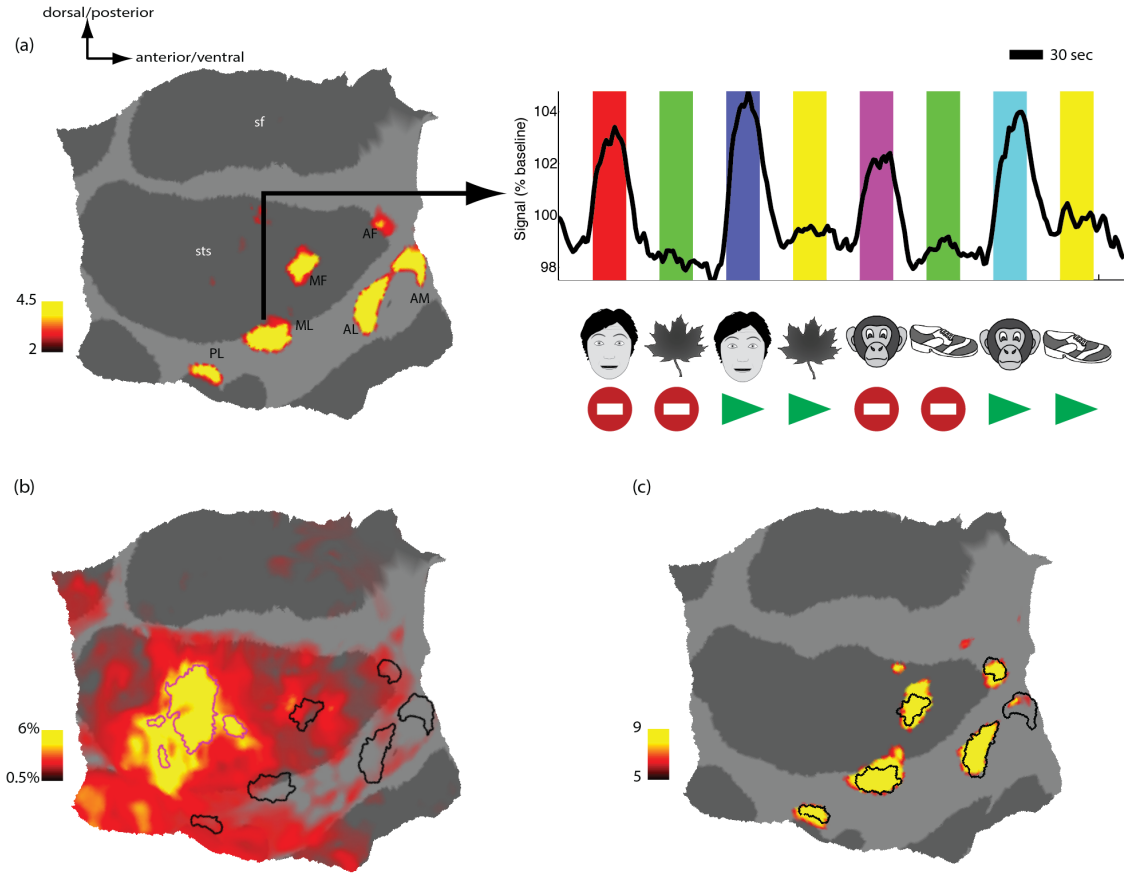

dorsal/posterior
(a)
anterior/ventral
sf
sts
AF
MF
AM
ML
AL
PL
4.5
2
30 sec
Signal (% baseline)
104
102
100
98
(b)
6%
0.5%
(c)
9
5

Table 2.1

Statistics of F-tests for main effects and interaction in macaque and human ROIs. For each contrast, F-value, degrees of freedom (DOF) and p-value (as negative common logarithm) are shown (see Materials and Methods for details on statistical tests).

Macaque

ROI		AM	AL	AF	ML	MF	PL
Contrast							
Shape category	**F**	195.8	460.4	330.7	548.3	199.7	227.8
	DOF	624	873	873	873	873	695
	P-val	38.2	81.6	62.2	93.8	40.3	44.0
Motion Condition	**F**	5.2	36.5	154.3	57.4	162.7	51.5
	DOF	695	873	873	873	873	695
	P-val	1.6	8.6	32.0	13.0	33.6	11.7
Interaction	**F**	8.0	15.3	0.2	6.9	0.1	2.4
	DOF	624	873	873	873	873	695
	P-val	2.32	4.0	0.2	2.1	0.1	0.9

Human

ROI		OFA	FFA	STS-FA
Contrast				
Shape category	**F**	126.8	178.5	155.9
	DOF	240	240	150
	P-val	22.4	29.1	24.5
Motion Condition	**F**	2.3	3.1	47.9
	DOF	240	240	150
	P-val	0.9	3.8	55.9
Interaction	**F**	2.3	0.6	12.0
	DOF	240	240	150
	P-val	0.9	0.4	3.2

We calculated the responses to static and moving faces and objects in a group analysis (Figure 2.2*a*) and identified the separate contributions of the factors: shape category (face vs. object), motion (moving vs. static), and their interaction (Figure 2.2*b*) in a two-way ANOVA (details in Materials and Methods, statistics in table 2.1). The main effect of shape category confirmed face selectivity in all face patches (Figure 2.2*b*; Table 2.1). The main effect of motion was strongest in MF and AF, weaker in posterior face patch PL, ML, and AL, and insignificant in AM (Figure 2.2*b*; Table 2.1). A subset of face patches, AL and AM, exhibited a significant interaction of motion with shape category (Figure 2.2*b*; Table 2.1). Thus, in face patches in the fundus of the STS (MF, AF), responses can be understood as a linear superposition of face selectivity and general motion sensitivity, whereas in patches furthest away from the fundus of the STS (AL and AM), the impact of stimulus motion is weaker and partially selective for facial motion.

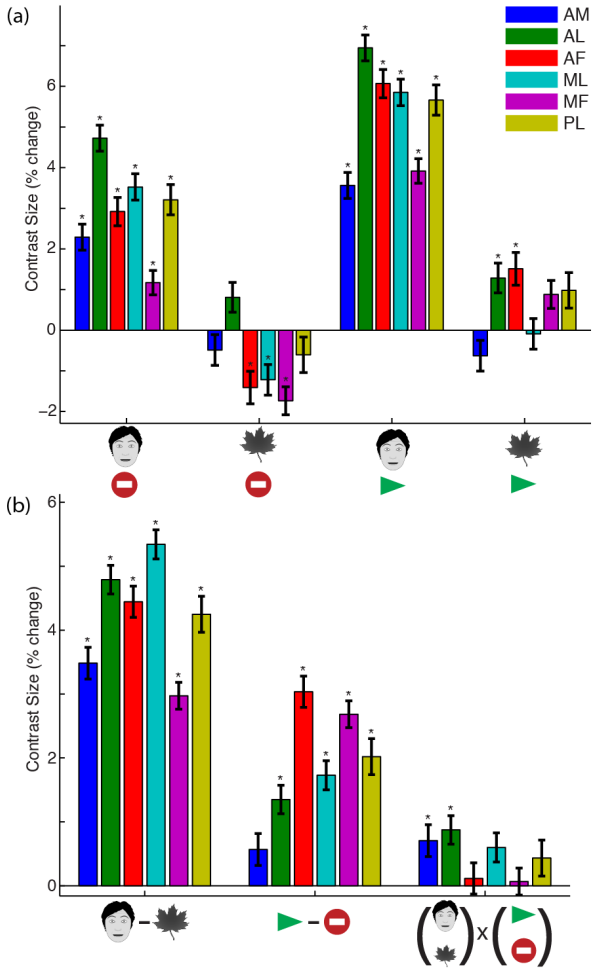


Figure 2.2

Responses to moving and static faces and objects in the macaque face patch system

(a) Group analysis responses to (from left to right) static faces, static non-face objects, moving faces and moving non-face objects in percent signal change from scrambled stimuli baseline in temporal lobe face patches. Asterisks mark significant differences from zero ($p<0.05$, Bonferroni-corrected for comparisons on multiple ROIs).

(b) Group analysis of (from left to right) the main effects of shape category, motion condition and their interaction. Asterisks as in (a).

The pattern of motion sensitivity along the fundus of the STS suggests a functional specialization of face patches by anatomical location in the STS. We tested this for two pairs of patches (ML vs. MF, and AL vs. AF) that differ in their location with respect to the fundus of the STS but are positioned at similar anterior–posterior positions along the STS. We performed a three-way ANOVA with ROI, motion and shape category. For the ML versus MF comparison, a significant two-way interaction between shape category and ROI ($F(1,1765) = 47.7$ $p < 10^{-11}$), and between motion and ROI ($F(1,1765) = 7.74$, $p < 10^{-2}$) indicated that the middle face patches differ in the strength of their selectivity to shape category (stronger in ML) and motion (stronger in MF). Nevertheless, a two-way interaction between motion ($F(1,1765) = 3.73$, $p = 0.053$) and shape category as well as a three-way interaction of ROI, motion, and shape category ($F(1,1765) = 2.42$, $p = 0.12$) did not reach significance. For the AL versus AF comparison, the two-way interaction between shape category and ROI was not significant ($F(1,1765) = 0.89$, $p < 0.35$), but the motion versus ROI interaction was ($F(1,1765) = 21.0$, $p < 10^{-5}$), with AF being more strongly selective for motion. For this pair, there was a significant two-way interaction between motion and shape category ($F(1,1765) = 7.18$, $p < 10^{-22}$) as well as a three-way interaction of ROI, shape category, and motion ($F(1,1765) = 4.27$, $p < 0.05$). This suggests a specialization for facial motion in AL absent in AF. We also analyzed the presence of a motion condition by shape category interaction on individual monkeys, focusing on the areas on the lip of the STS (AL and ML) and the

fundus of the STS (AF and MF). The effect was significant on the lip patches of Monkey M1 ($F(1,446) = 5.1$, $p < 0.05$) and both the lip ($F(1,172) = 5.1$, $p < 10^{-3}$) and fundus ($F(1,172) = 5.1$, $p < 10^{-3}$) patches of Monkey M3. This indicates the strength of specialization for facial motion is subtle and not always apparent in single individuals. Together, these analyses show strong effects of shape category and motion condition according to the position of face patches with respect to the STS.

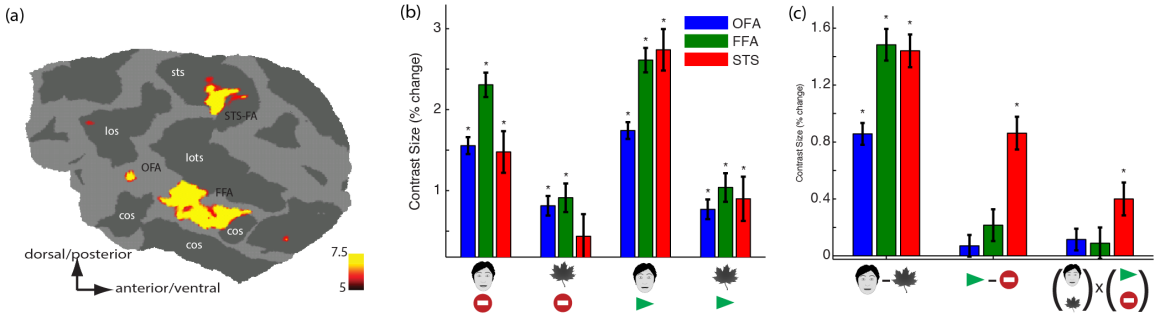


Figure 2.3

Responses to moving and static faces and objects in human face areas.

(a) Face selective regions on the flattened surface of the right hemisphere of a representative human subject. Color bar as in Fig. 2.1a. OFA, occipital face area; FFA, fusiform face area; STS-FA superior temporal sulcus face area. Sulci: los, lateral occipital; sts, superior temporal; lots, lateral occipito-temporal; cos, collateral.

(b) Group analysis of responses. Conventions as in Fig. 2.2a.

(c) Group analysis of main effects and interaction. Conventions as in Fig. 2.2b.

In humans, we identified face-selective regions (Figure 2.3*a*) FFA (left in 5 of 6 subjects, right in 6 of 6 subjects), OFA (left in 5 of 6 subjects, right in 6 of 6 subjects), and STS-FA (left in 1 of 6 subjects, right in 6 of 6 subjects). In Experiment 2, we calculated separate responses to all stimulus conditions (Figure 2.3*b*) and calculated the main effect of shape category, motion, and their interaction (Figure 2.3*c*). OFA and FFA activations were not significantly modulated by general motion (Figure 2.3*c*; Table 2.1). In contrast, the STS-FA exhibited significant modulation by motion and an interaction between shape category and motion (Figure 3.3*b*; Table 2.1). Underscoring the impact of face-specific motion on STS-FA responses, in 5 of 6 subjects the left STS-FA was found contrasting responses to moving faces versus moving non-face objects, but not contrasting static faces versus static objects. While these results mainly reproduce known specializations for facial motion in humans, they reveal new ones in macaques.

The implications of the present results for our understanding of face processing in both species need to be carefully considered. In brief, we could reproduce a known specialization for facial motion in the human face processing system that is not exactly matched by any face-processing area in the macaque. The results, then, don't lend themselves to a direct equalization between monkey and Hunan faces areas. In Chapter 4, we discuss the alternative scenarios that are compatible with these observations.

Chapter 3

Maps for spatial selection in the ventral and dorsal streams: A parietal priority map and a ventral attention spotlight?

The motivation for this set of experiments is to directly test the idea that areas that control spatial attention compute maps that encode a single priority signal, agnostic to its relationship with covert attention or eye movements, by summing disparate sources (Figure 3.1a). This scheme makes a number of predictions. First, we expect behaviorally relevant locations to be highlighted by cell responses, even in the physical absence of a visual stimulus. In addition, when different spatial sources that are important in a given task overlap in space, we expect their combined response to be the sum of those they would separately elicit. A third prediction is that the summed contribution of these sources constitutes one single priority signal that downstream areas can use to guide attention or behavioral responses when appropriate. A corollary of this is that, when attention and behavioral responses are spatially allowed to overlap, they cannot be simultaneously inferred from responses of cells in the priority map.

Figure 3.1

Conceptual motivation, brain areas involved and task design.

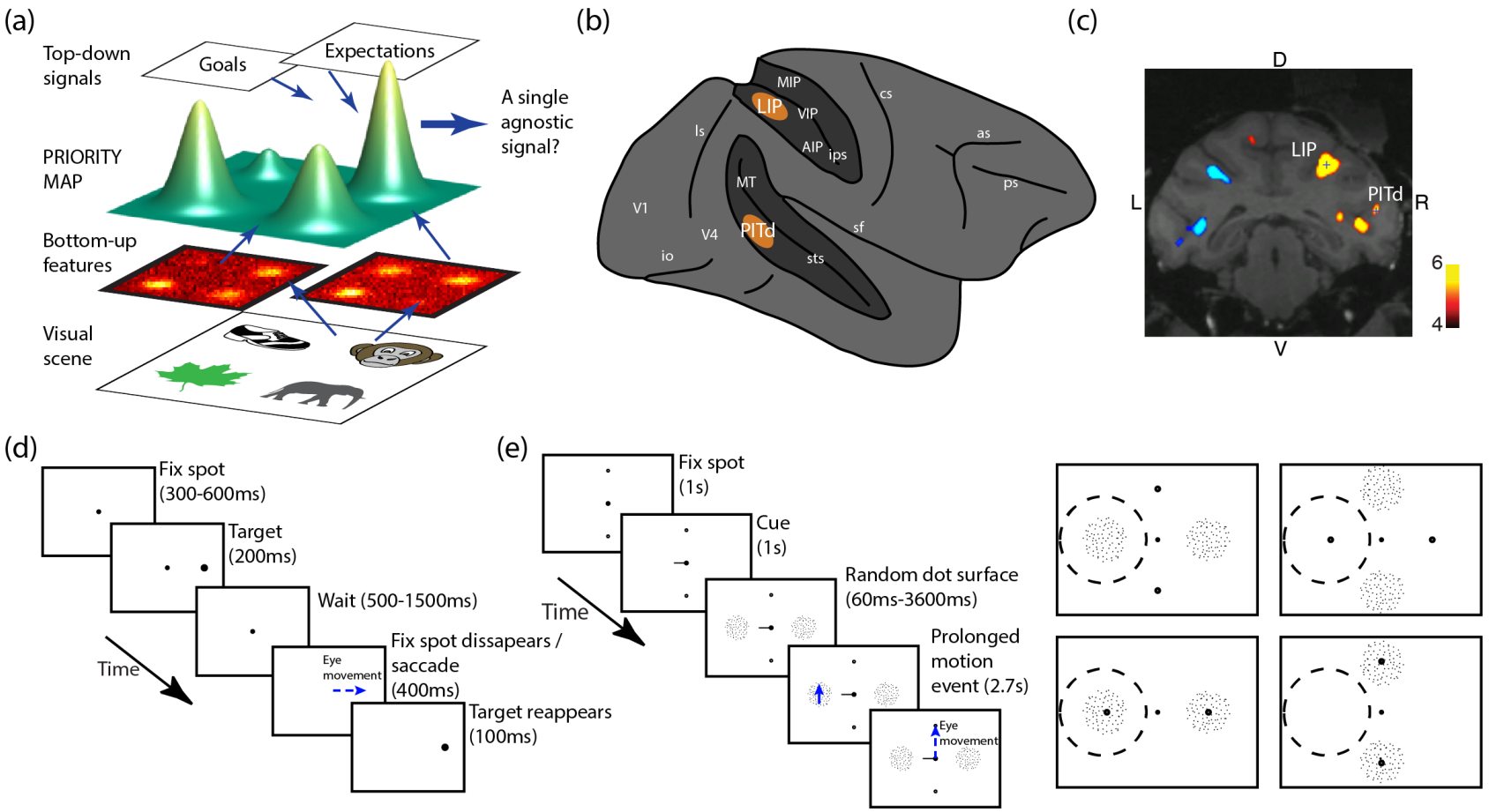
(a) Conceptual scheme of the priority map idea. An agnostic general-purpose priority signal is computed by summation of bottom-up signals from feature-coding areas and top down signals related to goals and expectations.

(b) Diagram of the macaque brain, showing the STS and IPS open, areas LIP and PITd highlighted in orange, together with other sulci and cortical areas for reference.

(c) Coronal MRI slice showing recoding locations in monkey M1 and a statistical map overlay (t-statistic, scale bar on right side) of covert attention effects from a previous study (Stemmann and Freiwald, 2015, in preparation), used to guide recordings.

(d) Schematics of the timing of events in a memory-guided saccade task. See Methods for details

(e) Schematics of the timing of events in a spatially cued discrimination task (left) and the four physical configurations of the stimuli relative to the receptive field of the cells being recorded (right), dashed-line circle. See Methods for details. The contrast of stimuli is inverted and size of central cue is exaggerated for illustration purposes.


(a)
Top-down signals
Goals
Expectations
PRIORITY MAP
A single agnostic signal?
Bottom-up features
Visual scene
(b)
MIP
LIP
VIP
cs
ls
AIP
ips
as
MT
ps
V1
V4
PITd
sf
io
sts
(c)
D
LIP
PITd
L
R
6
4
V
(d)
Fix spot (300-600ms)
Target (200ms)
Wait (500-1500ms)
Time
Eye movement
Fix spot dissapears / saccade (400ms)
Target reappears (100ms)
(e)
Fix spot (1s)
Cue (1s)
Random dot surface (60ms-3600ms)
Time
Prolonged motion event (2.7s)
Eye movement

As discussed in chapter 1, two possible alternatives to the priority map hypothesis could be “pure attention” and the “pure intention” scenarios. In the “pure intention” scenario, the population encodes a motor plan only. In this case, effects of covert attention, if observed, indicate the planning of an eye movement. What distinguishes this scenario from others it that, right before a saccade, then, there should be no signals indicating a different, covertly relevant, location. Conversely, in the “pure attention” case, the population represents covertly attended locations and the observation of different signals like eye movements occur only if there is a covert shift of attention before a saccade. In this scenario, as opposed to others, there should be no distinct saccadic signal, when the eye movement is made to an already attended location.

In order to test the three predictions of the priority map hypothesis, we performed extracellular electrophysiological recordings in areas PITd and LIP (Figure 3.1b), both of which have been proposed to control spatial attention, in two monkeys (M1 and M2) performing behavioral tasks designed for this purpose. Selection of recording locations was guided by fMRI results (see Methods and Figures 3.1c and S3.1) from a previous study with the same animals (Stemmann and Freiwald, 2015, in preparation).

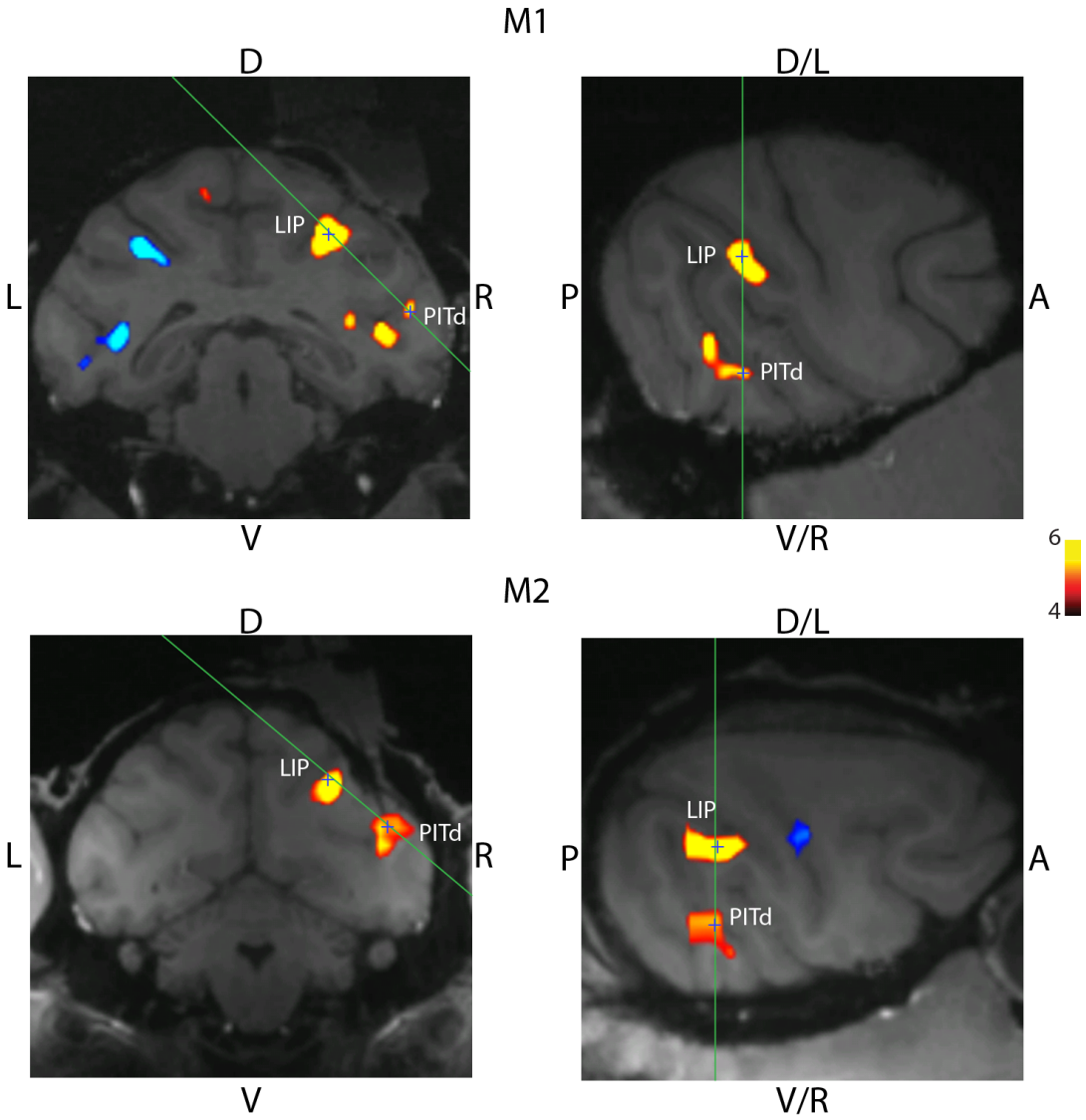
Two monkeys were trained to perform two behavioral tasks. The first one is a memory-guided saccade (MGS) task (see Methods and Figure 3.1d), which is used routinely in LIP recordings to confirm recording locations (e.g., Roitman and Shadlen, 2002) and similarly in FEF studies to classify cell subpopulations

(e.g., Gregoriou et al., 2012). Trials in the MGS began with a central fixation period followed by a peripheral visual target presentation. After extinction of the target, monkeys waited for a variable period of time (memory period, 400ms to 900ms) until the central fixation spot disappeared, which cued them to make a saccade to the remembered location. Existence of tuned neural activity to the location of the remembered target during the memory period of the MGS task is expected in cells encoding a priority map.

Figure S3.1

MR-guided electrophysiology.

Coronal and sagitto-horizontal planes showing recoding locations (marked with a blue cross) targeting PITd and LIP in monkeys M1 and M2. Green lines indicate the orientation of the related slice in the adjacent panel. Colored overlay is a statistical parametric map of covert attention effects used to guide recordings (t-statistic, scale bar on right side). PITd location in monkey M1 is 0mm anterior to the interaural line (+0, AP) axis, +23 on the medio-lateral (ML) axis and +18 along the dorso-ventral (DV) axis. LIP in monkey M1 was targeted at AP +0, ML +12, DV +28. PITd in monkey M2 was targeted at AP -1.5, ML +23, DV +21. LIP in monkey M2 was targeted at AP -2, ML 14, DV +29.


M1
D
D/L
LIP
PITd
L
R
P
A
V
V/R
6
4
M2
D
D/L
LIP
PITd
L
R
P
A
V
V/R

A second, spatially cued discrimination task was designed to study the simultaneous encoding of different spatial signals (Figure 3.1e). Trials started with a central fixation period after which a small central cue cued monkeys in the direction of a future relevant stimulus. After 1s, two moving-dot surfaces appeared on opposite sides of the visual field, the cued one being behaviorally relevant. A fraction of dots in each surface moved coherently in the same direction. For each surface, direction of motion changed independently every 60ms for a variable period of time (60 to 3600ms, independent for each surface) until a prolonged motion event (PME) during which the direction of motion remained the same. Direction of motion during the PME was in one of two opposite directions. Monkeys had to indicate the direction of motion of the cued surface, as soon as they made a decision, with an eye movement to one of the saccade targets.

The crucial component of the spatially cued discrimination task is the spatial position of task elements (moving-dot surfaces and saccade targets) relative to the receptive field of the cell being recorded. There were four unique physical stimulus arrangements (see Methods and Figure 3.1e, right panel) in which moving-dot surfaces and saccade targets would be positioned either inside or outside the receptive field. For each arrangement, there were four unique behavioral situations (two possible cued surfaces and two possible saccade directions), making a total of 16 unique conditions overall.

Task difficulty was regulated by adjusting values of motion coherence, to compensate variations due to changing eccentricity of the stimuli (determined by the receptive field of the neurons recorded) and training levels at each location (similar receptive fields were typically observed in consecutive days, over which behavioral performance would improve). Coherence values between 9 and 28% were used, to keep daily performance between 66 and 75% in trials in which the monkey made a response. Correct performance in the task required reaching a decision across several hundred milliseconds (mean and standard deviation across sessions of median reaction times: M1, PITd sessions: 720± 120ms; M1, LIP: 880± 130ms; M2, PITd: 780± 130ms; M2, LIP: 840± 140ms).

We report analyses using 85 PITd units and 66 LIP units, recording from all units that had a tuned visual response (see Methods). Since the main findings were consistent across monkeys, we pooled units from both for analyses. Individual results are reported in Supplementary Figures.

Memory guided saccade (MGS) task

We first studied neural responses during the MGS task with the main goal of determining if there was tuned activity during the memory period of the task in both areas. Persistent activity during the memory period of the MGS task is thought to provide a bridge between sensory input and motor output, suggesting that such populations can provide a window into simple forms of higher cognition

(Huk and Meister, 2012; Shadlen and Gold, 2004). We computed spike density functions from individual trials and compared the responses for trials in which the stimulus fell inside the RF to those in which it appeared opposite to it (see Methods). Since trial duration was variable, we studied the activity aligned on each of two key events: stimulus onset and saccade onset (see Methods). As a consequence, the number of trials contributing to the mean traces varies across time.

Time course of activity for typical units from both areas is shown in Figure 3.2 (left panels). Both units showed a brisk response after stimulus onset, followed by sustained tuned activity that persisted until saccade onset.

To quantify these effects at the population level, responses were z-scored as a way to factor out the overall variability in firing rates across different units. Tuning during the memory period was computed by calculating response differences between trials with target inside the RF and opposite to it, using a time window from 600 to 400ms before saccade onset. Large fractions of cells in both areas displayed significantly tuned activity during the memory period of the task (Figure 3.2, center panels; Figure S3.2, left panels). To determine if there were subpopulations with a distinct premotor component, we followed the approach of Gregoriou and colleagues (2012), and computed the difference between activity in the memory period and a presaccadic period (200 to 0ms before saccade). We did not find evidence of distinct subpopulations, but rather a continuous distribution, similar in both areas, with a small fraction of cells

showing a significant effect (Figures 3.2 and S3.2, right panels). These results suggest both PITd and LIP treat attended/remembered locations similarly, are compatible with the priority map hypothesis in both areas and don't reveal distinct visual and premotor subpopulations.

Figure 3.2

Neural responses during memory-guided saccade (MGS) task.

Left: Spike density functions of example PITd (top) and LIP (bottom) units, with timing relative to stimulus onset (left half, duration indicated by black rectangle) and saccade onset (right half) in trials where target was inside the receptive field (blue trace) and opposite to it (green trace). Error bars indicate 95% confidence intervals. Dashed vertical lines represent alignment time.

Center: Histogram of spatial tuning of z-scored activity in the memory period of the MGS task in PITd (red, top) and LIP (blue, bottom) combined across monkeys. Filled segments show statistically significant units (two-tailed t-test, $p<0.05$).

Right: Histogram of strength of presaccadic activity (z-scored) increase in the MGS task in PITd (red, top) and LIP (blue, bottom). Filled segments show statistically significant units (two-tailed t-test, $p<0.05$).

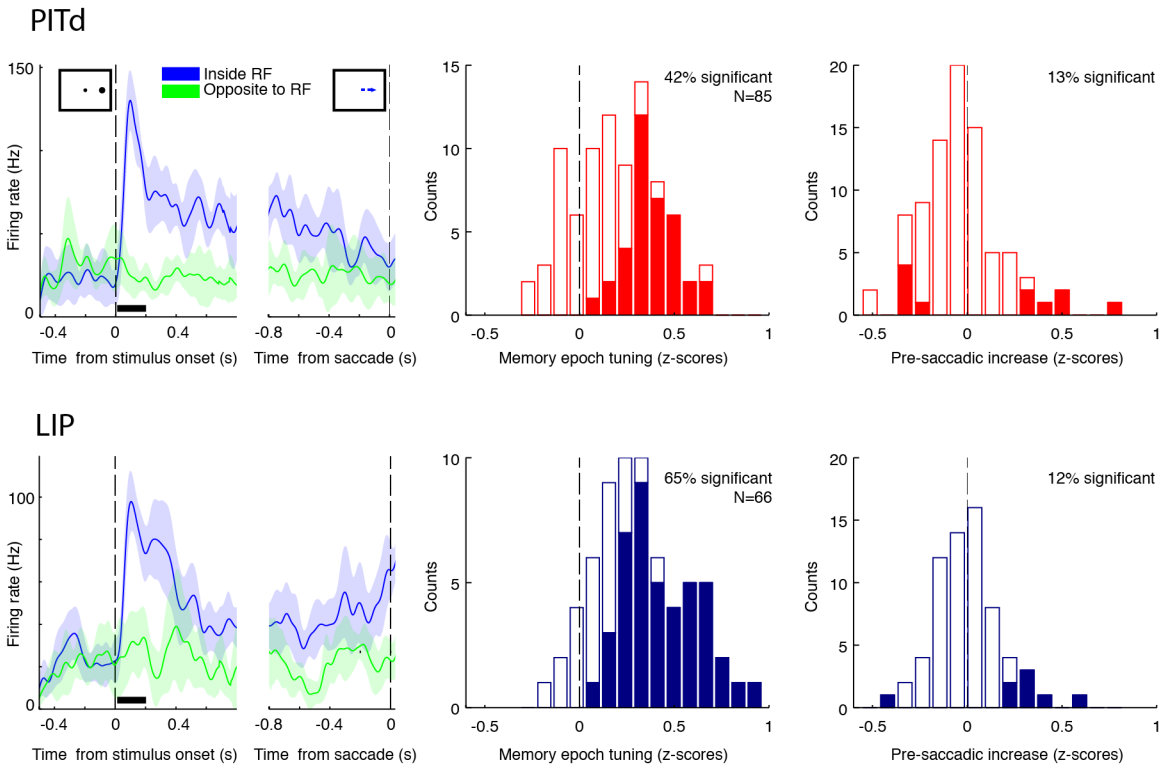
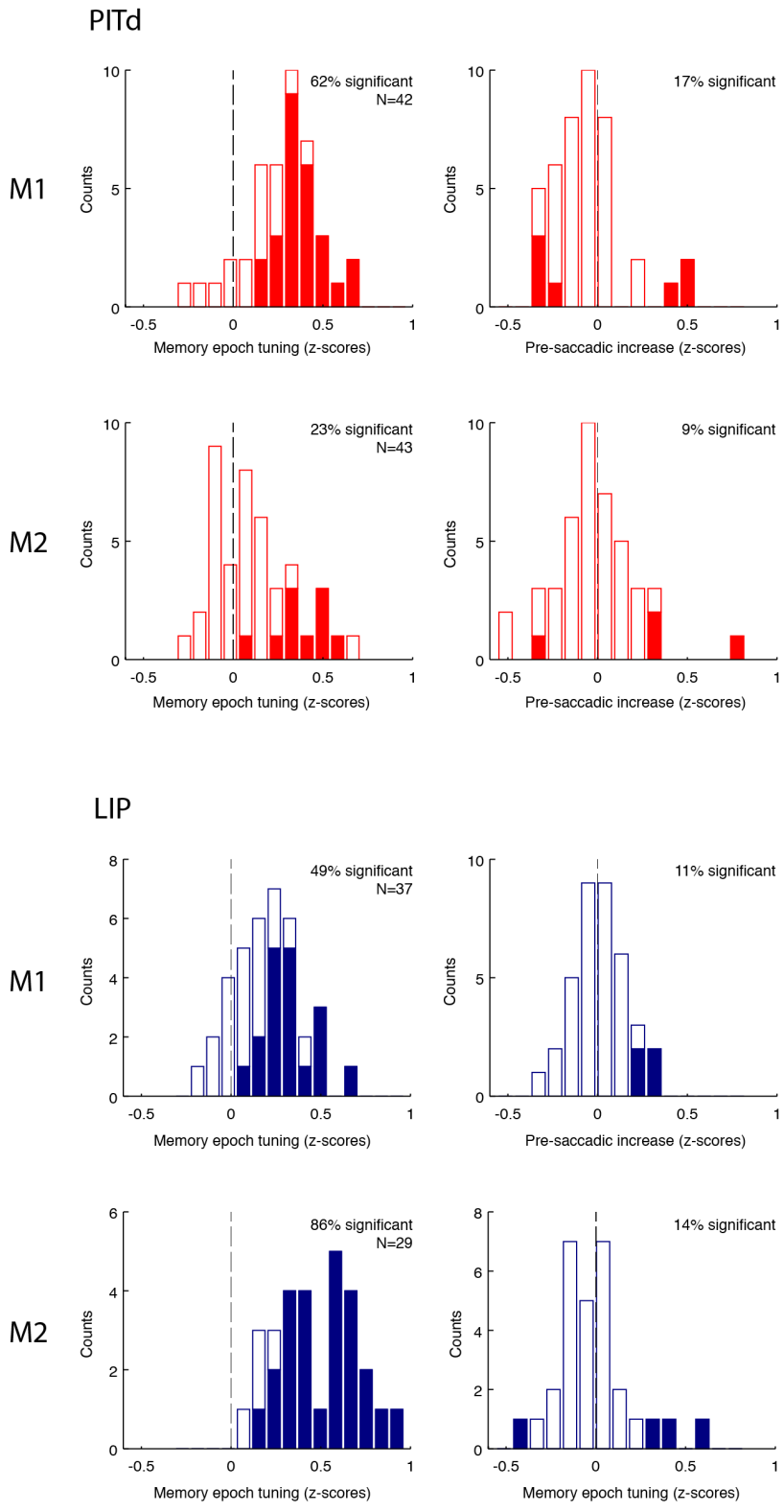

PITd
Inside RF
Opposite to RF
Firing rate (Hz)
Time from stimulus onset (s)
Time from saccade (s)
Counts
42% significant
N=85
Memory epoch tuning (z-scores)
13% significant
Pre-saccadic increase (z-scores)
LIP
65% significant
N=66
12% significant

Figure S3.2

Neural responses during memory-guided saccade (MGS) task in individual monkey.

Left column: Histogram of z-scored spatial tuning of activity in the memory period of the MGS task in PITd (red, top two plots) and LIP (blue, bottom two plots). Filled segments show statistically significant units (two-tailed t-test, $p<0.05$).

Right column: Histogram of strength of presaccadic z-scored activity increase in the MGS task in PITd (red, top two plots) and LIP (blue, bottom two plots). Filled segments show statistically significant units (two-tailed t-test, $p<0.05$).


PITd
M1
62% significant
N=42
17% significant
Counts
Memory epoch tuning (z-scores)
Pre-saccadic increase (z-scores)
M2
23% significant
N=43
9% significant
LIP
M1
49% significant
N=37
11% significant
M2
86% significant
N=29
14% significant

Neural responses during spatial spatially cued discrimination task: Encoding of task variables

The explicit goal of the experimental design was to study the representation of different spatial signals and how their effects when studied in isolation relate to those when studied in combination. A first step, then, is to compare the activity for each unit in the different task conditions. We computed spike density functions (SDFs) from individual trials and compared the mean time courses in each condition. Since trial duration was variable, as in the MGS task, we studied the activity aligned to each of two key events (moving-dot surface onset and saccade onset), excluding activity on time points close to the other event (see Methods).

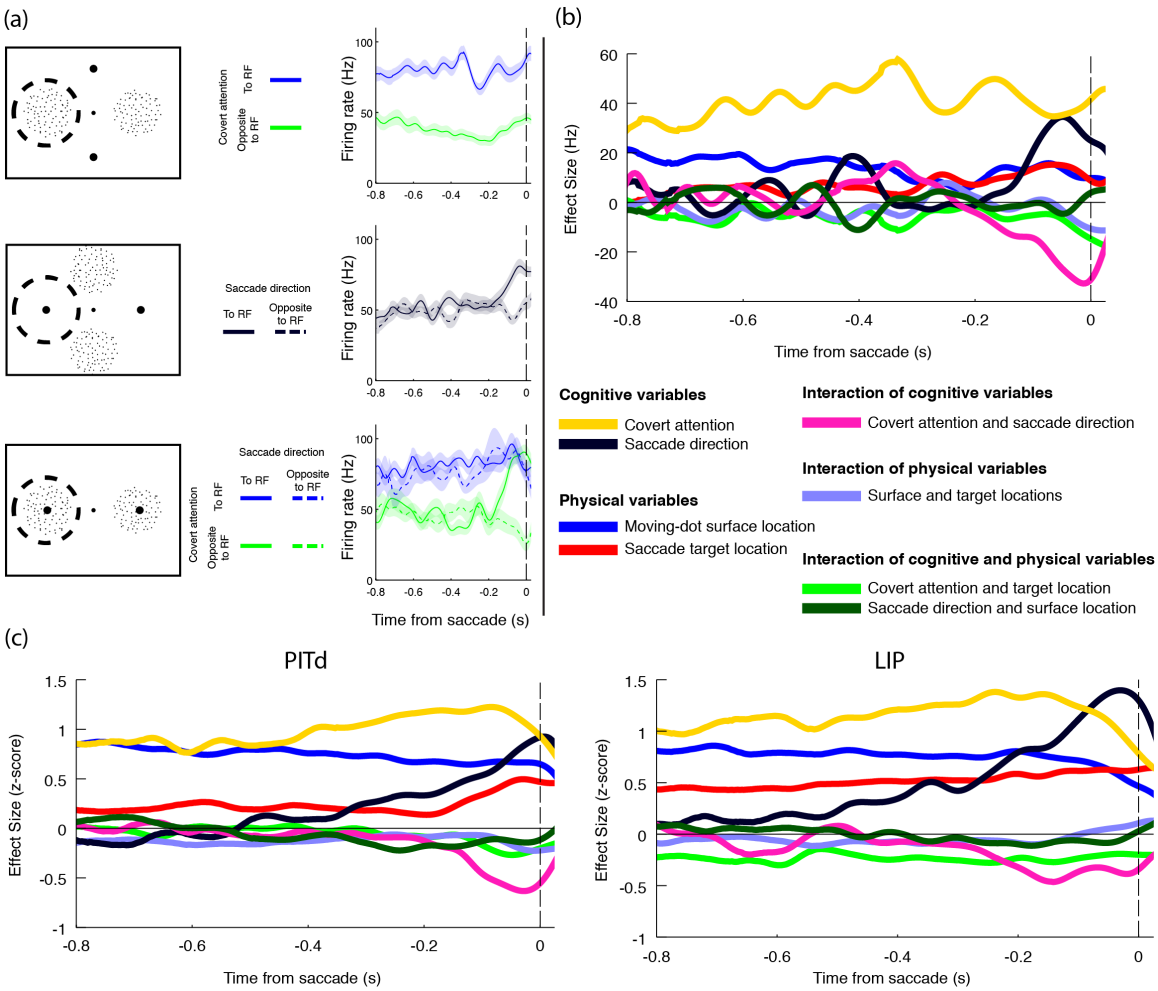


Figure 3.3

Neural responses during the spatially cued discrimination task.

(a) Spike density function, with time aligned to saccade onset, of an example PITd unit when inside the receptive field there are only dot surfaces (top, covert attention indicated by color), or only a saccade target (center, saccade direction indicated by continuous or dashed line) or both (bottom, covert attention indicated by color, saccade direction indicated by continuous or dashed line).

(b) Results of GLM analysis applied to the unit shown in (a), with time aligned to saccade onset. Colors indicate task variables and their interactions as indicated in legend.

(c) Mean z-scored results of GLM analysis, with time aligned to saccade onset, in areas PITd and LIP of both monkeys combined. Color scheme as in (b).

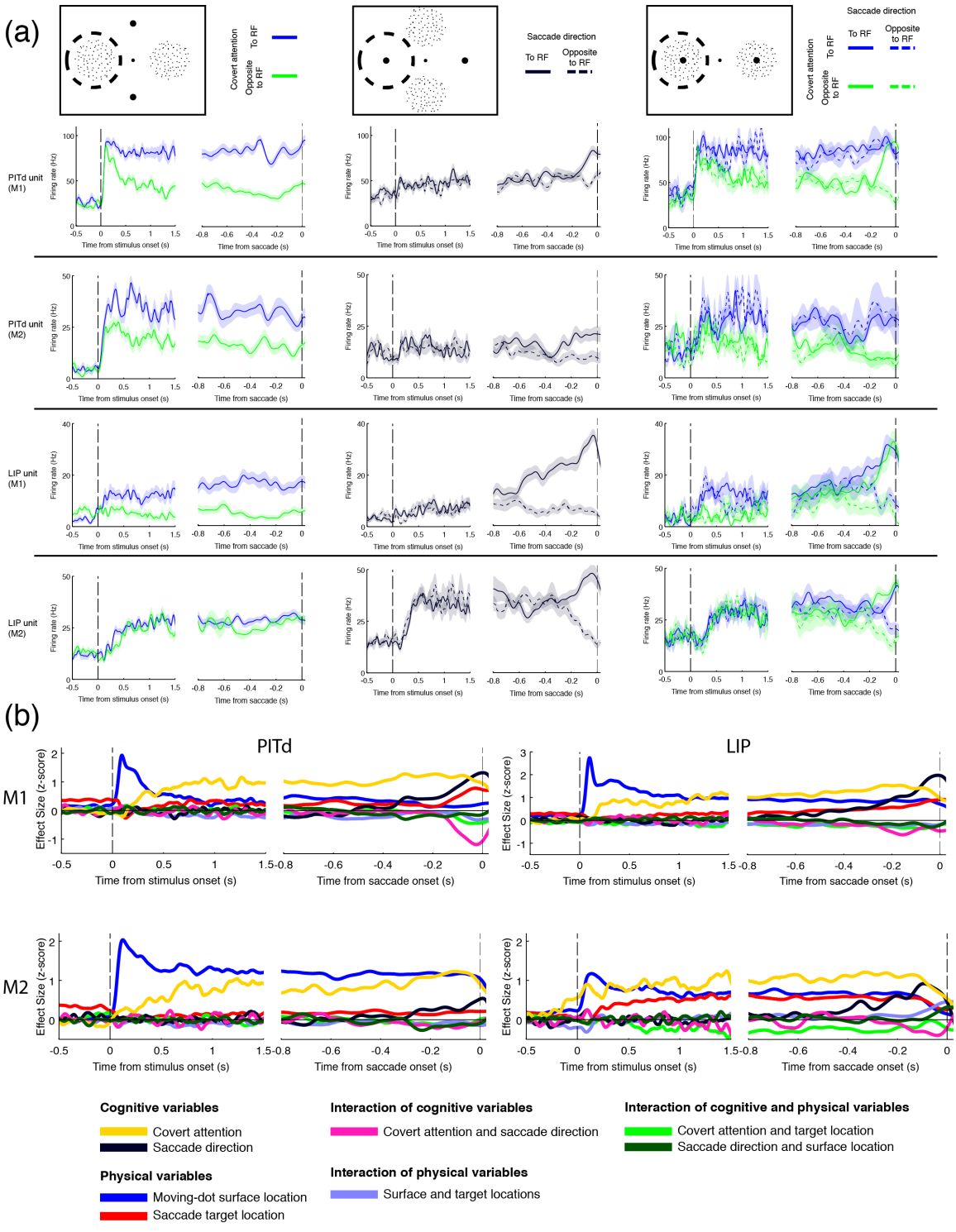
Figure S3.3

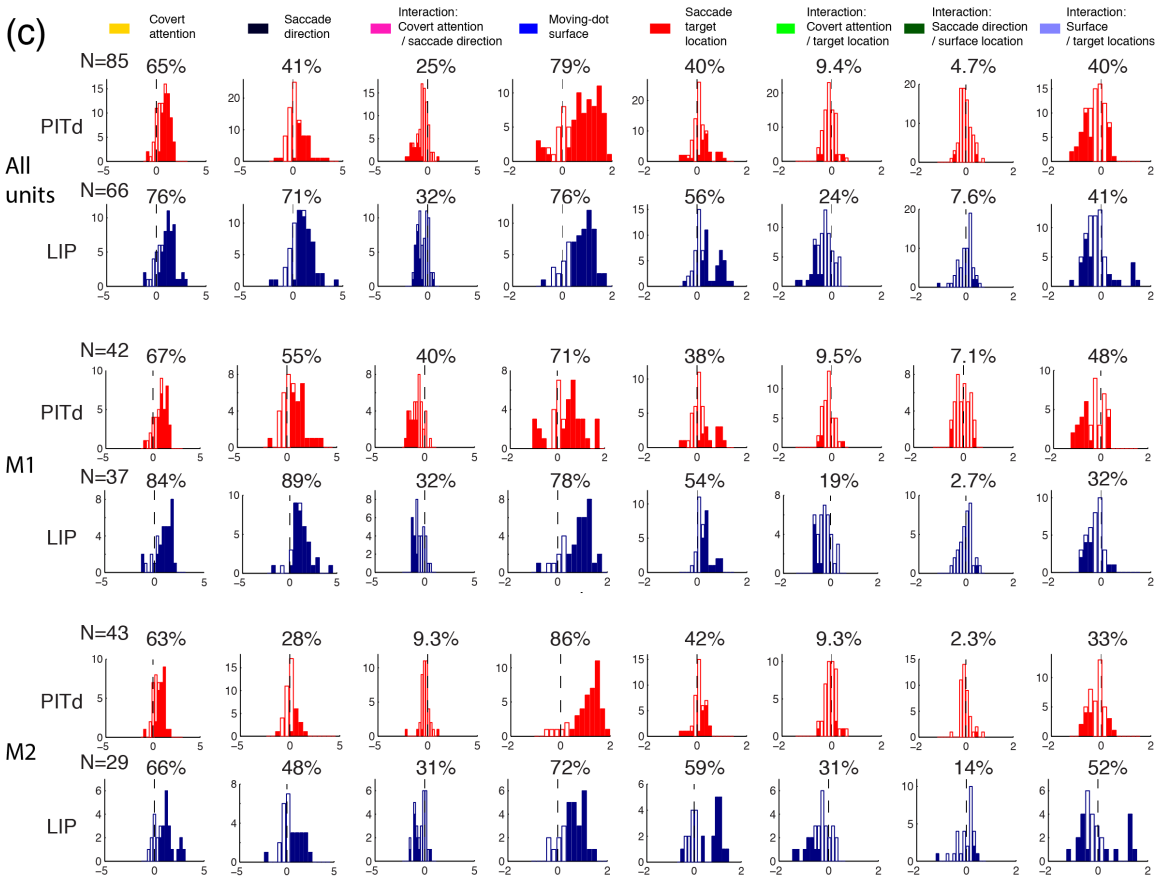
Neural responses during the spatially cued discrimination task, continued.

(a) Spike density function of example units (top row is the same unit as Figure 3a), aligned on dot surface onset (left half) and saccade onset (right half). Conventions like in Figure 3.3a.

(b) Mean z-scored results of GLM analysis, with time aligned to dot surface onset (left half) and saccade onset (right half), in areas PITd and LIP in individual monkeys. Color scheme as in Figure 3.3b.

(c) Distribution of effect sizes (z-scored) of task variables and their interactions (see legend on top, color code from GLM time courses provided for reference) in areas PITd (red plots) and LIP (blue plots) in both monkeys combined (top two rows) and in individual monkeys M1 (third and fourth row) and M2 (bottom two rows). Filled segments indicate statistically significant units (two-tailed t-test, $p<0.05$). Percent numbers on top of each plot denote fraction of statistically significant units. Units on all x-axes are z-scores and those on all y-axes are unit counts.


(a)
Covert attention
To RF
Opposite to RF
Saccade direction
To RF
Opposite to RF
PITd unit (M1)
PITd unit (M2)
LIP unit (M1)
LIP unit (M2)
Firing rate (Hz)
Time from stimulus onset (s)
Time from saccade (s)
(b)
PITd
LIP
M1
M2
Effect Size (z-score)
Time from stimulus onset (s)
Time from saccade onset (s)
Cognitive variables
Covert attention
Saccade direction
Physical variables
Moving-dot surface location
Saccade target location
Interaction of cognitive variables
Covert attention and saccade direction
Interaction of physical variables
Surface and target locations
Interaction of cognitive and physical variables
Covert attention and target location
Saccade direction and surface location


(c)
Covert attention
Saccade direction
Interaction: Covert attention / saccade direction
Moving-dot surface
Saccade target location
Interaction: Covert attention / target location
Interaction: Saccade direction / surface location
Interaction: Surface / target locations
All units
PITd
N=85
65%
41%
25%
79%
40%
9.4%
4.7%
40%
LIP
N=66
76%
71%
32%
76%
56%
24%
7.6%
41%
M1
PITd
N=42
67%
55%
40%
71%
38%
9.5%
7.1%
48%
LIP
N=37
84%
89%
32%
78%
54%
19%
2.7%
32%
M2
PITd
N=43
63%
28%
9.3%
86%
42%
9.3%
2.3%
33%
LIP
N=29
66%
48%
31%
72%
59%
31%
14%
52%

Our analyses focused on time courses aligned to saccade onset, since it is around that event that all task variables become most relevant simultaneously. Time courses of an example PITd unit are shown in Figure 3.3a. When only a moving-dot surface was placed inside the receptive field (Figure 3.3a, top panel), responses were heavily modulated by covert attention until the end of the trial. When only a saccade target was inside the receptive field (Figure 3.3a, middle panel), saccade direction selectivity emerged in the last 200ms before the end of the trial. When both a moving-dot surface and a saccade target were inside the receptive field (Figure 3.3a, bottom panel), covert attention tuning and saccade selectivity interacted, giving rise to a more complex pattern. There, if covert attention was directed toward the receptive field, firing rate stayed at a high level throughout the trial and there was no saccade selective signal. In contrast, when covert attention was directed away from the receptive field, activity was low until the last 200ms of the trial, when strong saccadic selectivity emerged. In this example unit, then, the effect of cognitive variables did not linearly add, and covert attention overrode the effects of presaccadic tuning. In consequence, by the end of the trial it is not possible to univocally determine the behavioral context from the neural activity. This particular cell is consistent with the “pure attention” scenario described at the beginning of this chapter: the effect of saccades can be observed only when it is accompanied by a covert shift of attention.

The effect of task variables across cells was heterogeneous. To illustrate the richness of neural repertoires, we provide more examples in Figure S3a,

where time courses are shown aligned both on the onset of moving-dot surfaces and saccades. In order to quantitatively test the prediction from the priority map proposal, data from all trials in each cell was modeled with a multiple linear regression (see Methods). This makes it possible to estimate the effect of each task variable and its interaction with every other. Task variables of most interest for our purposes describe the cognitive state of the monkey (covertly attended surface and selected response) and their interaction, but we computed the effect of "physical" factors (the presence of a moving-dot surface or a saccade target inside the receptive field) as well as their interaction with the cognitive variables and between themselves. Under the priority map prediction, interactions would be negligible.

Results of the linear-regression analysis in the example PITd unit (Fig. 3.3a) are shown in Figure 3.3b. The strongest effect was that of covert attention (estimated in trials where only moving-dot surfaces where in the receptive field), staying at a roughly constant level throughout the trial. The effect of saccade direction appeared in the final 200ms, reaching a similar effect size than that of covert attention. The interaction of covert attention and saccade direction had a comparable magnitude but opposite sign. This indicates that the two factors combine sub-additively and that the size of this nonlinearity was such that pre-saccadic tuning was completely abolished when covert attention was directed toward the receptive field. Compared to these two cognitive factors, the physical

effect of surface and saccade target presence inside the RF impacted the cell's response less; all other interactions were small.

To quantify effects at the population level in each area, we normalized the activity of each unit (conversion to z-scores) to factor out variability in overall firing rates and computed the average effect of task variables and their interactions using all cells. Results are shown in Figure 3.3c and S3.3c (individual monkeys, aligned on both stimulus and saccade onset). The patterns in PITd and LIP resembled that of the example cell and were strikingly similar to each other. The main effect of covert attention to the moving-dot surfaces inside the receptive field stayed constant throughout the trial, while tuning to saccade direction built up slowly until the end of the trial. The interaction of covert attention and saccade direction was somewhat bigger in PITd, especially relative to the size of saccadic tuning, which was more prominent in LIP. Other differences include a stronger effect of the saccade target presence inside the receptive field in LIP and a tendency in LIP to have negative interactions between covert attention and target presence, indicating that covert attention effects in LIP are reduced when potential saccade targets overlap with the moving-dot surfaces. We computed the strength and statistical significance of all these effects in fixed time windows (from 200 to 0ms from saccade onset in the case of saccade direction effect and its interactions, and 800 to 500ms for all other effects and interactions) and computed distributions for both areas in combined and individual monkeys (Figure S3.3c), which are consistent with the described

pattern. The percentage of units showing significant effects and interaction is roughly similar in both areas, as well as the range of the effect sizes. An important difference is the fraction of units with significant pre-saccadic tuning, which is almost double in LIP than PITd. In addition, there are signs of a bimodal distribution of responses to saccade target location, most notoriously in LIP, suggesting differential processing within the population. Furthermore, between 20% and 30% of LIP cells show a significant interaction between effects of covert attention and saccade presence. This indicates effects of covert attention are reduced when surfaces and saccade targets overlap in those units, which is a deviation from the scheme of simple sums

A word of caution is necessary when interpreting the results from the linear regression. A first caveat is that linear regression does not incorporate ceiling effects that might take place due to biophysical limitations of firing mechanisms. In that sense, linear regression is a study of the resulting representation rather as opposed to its underlying mechanisms. Second, linear regression doesn't take into account the fact that spiking activity follows a Poisson statistic, with the variance increasing together with firing rates.. A future step then is to use a generalized linear model that incorporates this fact explicitly. Third, just because a certain interaction is statistically significant, it doesn't mean that it has functional consequence. In that regard, the size of interactions is more important and it has to be compared with the magnitude of single effects to see

the degree to which interactions enhance or abolish the effect of certain task variables.

Taken together, these analyses show remarkable similarities in responses to task variables in PITd and LIP with both areas displaying signals of covert attention and response selection. In both areas, there are interactions between task variables, indicative of departures from the priority map proposal. From these results alone, it is unclear if the interactions affect the information present in the population about spatial cognitive signals when these overlap in space nor whether the population responses to separate signals can be used to interpret results of the same signals combined. Subsequent analyses address these issues directly.

Figure 3.4

Decoding of cognitive task variables in the spatially cued discrimination task.
Decoding accuracy, at time points aligned to saccade onset, combining cells from both monkeys from PITd (red) and LIP (blue).
(a) Decoding of covert attention location with only a dot-surface inside the receptive field.
(b) Decoding of covert attention location with both a dot-surface and a saccade target inside the receptive field.
(c) Decoding of saccade direction with only a saccade target inside the receptive field.
(d) Decoding of saccade direction location with both a dot-surface and a saccade target inside the receptive field.
Shading indicates standard deviation across resampling pseudo-populations. Colored horizontal segments denote time points when performance differs significantly from chance levels (permutation test, $p<0.01$) for PITd (red) and LIP (blue).

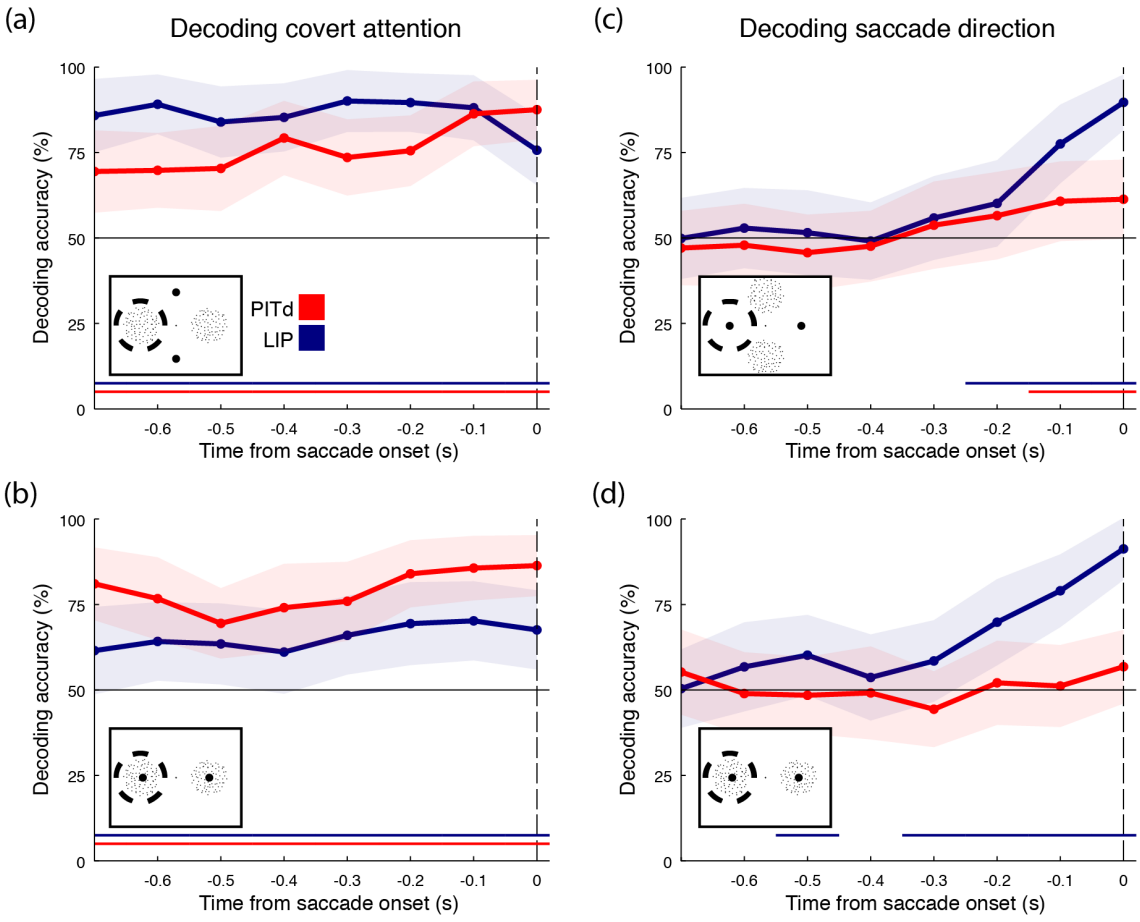
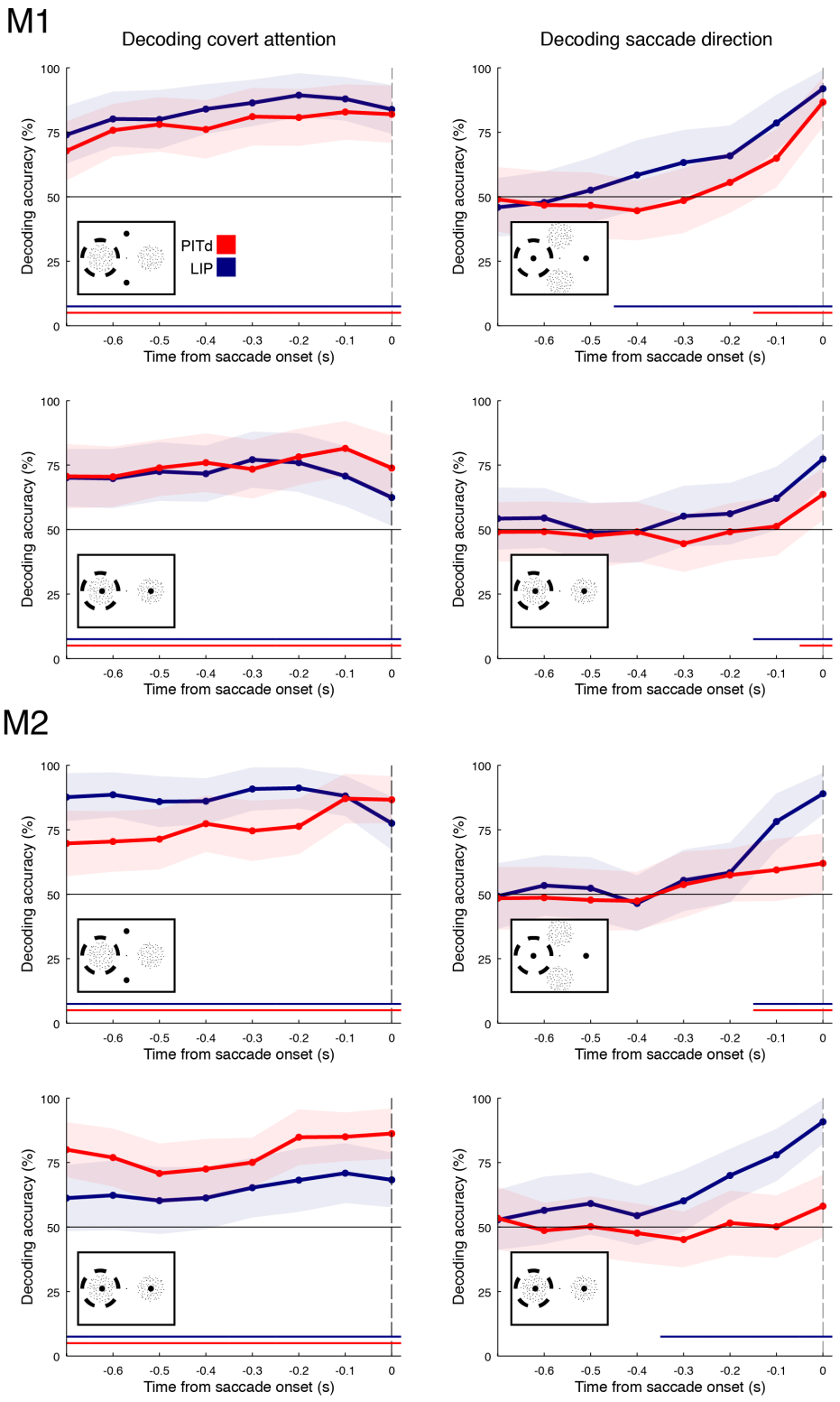

(a)
Decoding covert attention
(c)
Decoding saccade direction
(b)
(d)
Decoding accuracy (%)
Time from saccade onset (s)
PITd
LIP
100
75
50
25
0
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0

Figure S3.4

Decoding of cognitive task variables in the spatially cued discrimination task in individual monkeys.

Decoding accuracy, at time points aligned to saccade onset, in monkeys M1 (top) and M2 (bottom) from PITd (red) and LIP (blue) to decode covert attention (left) and saccade direction (right) when only one stimulus is inside the receptive field (odd rows, see inset diagram) or both are (even rows, see inset diagram). Shading indicates standard deviation across resampling pseudo-populations. Colored horizontal segments denote time points when performance differs significantly from chance levels (permutation test, $p<0.01$) for PITd (red) and LIP (blue).


M1
Decoding covert attention
Decoding saccade direction
PITd
LIP
Decoding accuracy (%)
Time from saccade onset (s)
M2

Neural responses during the spatially cued discrimination task: Decoding of task variables

While some information about task variables can be inferred from inspection of individual cells or even mean responses in the population, there is no direct correspondence between them and the information distributed in the population as a whole. To directly address this question, we used a decoding approach to test whether information is contained about covert attention and saccades when they are present in isolation and when they can overlap in space in both areas. We used linear support vector machines (SVM) (Cortes and Vapnik, 1995) to classify task conditions (see Methods), which, while simple, can be interpreted as performing computations of a downstream neuron and are particularly suitable for high-dimensional representations (Rigotti et al., 2013). In that sense, it is said that linear SVMs are sensitive to information that is explicitly represented in the population. Certainly, decoding results obtained this way provide a lower bound for the amount of information available at the population level, which could in principle be extracted using more elaborate devices. We investigated the information present in collections of cells that correspond to a given location in the priority map, by labeling all task conditions in a system of coordinates relative to each unit's receptive field, which were then binned in fixed 100ms periods and pooled together to form a pseudo-population. Pseudo-populations were constructed with an equalized number of units (20), and the results reported

correspond to the average of many (200) such populations, which are assembled by randomly sampling with replacement from the original pool of cells. This has the advantage of ensuring that the results do not depend on the inclusion of any given cell (which is typically absent in more than 70% of the subpopulations when units from both monkeys are combined) and providing an estimate of the variability within the total population.

Decoding results of covert attention and saccades for both areas are shown in Figures 3.4 and S3.4 (individual monkeys), relative to the time of saccade onset. Decoding accuracy of covert attention in the case when only moving-dot surfaces are present in the receptive field (Figure 3.4a) was significantly above chance levels for all time points in both areas. In the condition when saccade targets were also positioned inside the receptive field (Figure 3.4b), decoding accuracy was also above chance level for all time points in both areas. Interestingly, decoding accuracy seemed to become somewhat reduced in LIP and not in PITd, relative to the condition with surfaces only in the receptive field. This observation is consistent with the interaction between effects of covert attention and position of saccade targets observed above (Figure S3.3c). The fact that covert attention location is explicitly present until the end of the trial in both areas, even while a saccade while is executed to an independent location, is critical because it enables us to discard a “pure intention” scenario in both areas. Decoding accuracy of saccade direction in trials with only saccade targets in the receptive field (Figure 3.4c) emerged gradually towards the end of the trial in both

areas. Accuracy appears to be smaller in PITd than LIP, mostly because effects are weaker in monkey M2 (Figure S3.4). In trials where moving-dot surfaces were also positioned inside the receptive field (Figure 3.4d), decoding of saccade direction was largely unaffected in LIP but did not differ significantly from chance levels in PITd (except for the single last time point in monkey M1).

The decoding results, then, suggest a difference between both areas. Information about both covert attention and saccades was present and robust in LIP, even on the same set of trials. This is not a prediction of the priority map proposal, which interprets the linear sum of disparate sources by neurons as computing a single combined quantity. It is not immediately clear from the decoding analyses if this simultaneous information is a consequence of the additivity of signals or it is just possible in spite of it. However, this discards both a “pure intention” and a “pure attention scenario” In PITd, information about saccade direction and covert attention existed separately, but decoding of saccades was severely affected by the presence of moving-dot surfaces. Again, the analyses cannot discern if this impairment was caused by a reduction of neural signal-to-noise ratios or because a true “pure attention” scenario is the case. To answer these questions, it is necessary to inspect neural dynamics in more than one dimension.

Neural responses during the spatially cued discrimination task: dynamics in task-relevant dimensions

In our analysis of how variables are encoded, interactions between task variables were found, but it was not clear if they were important enough to invalidate the priority map proposal. Conversely, we found information about covert attention and task variables was simultaneously present in LIP but not in PITd, though it was not immediately obvious how that finding related to the proposed encoding scheme. In order to bridge the linear regression and decoding results, we used targeted dimensionality reduction (Mante et al., 2013) to understand the dynamics along axes in population space that explain the variance due to task variables (see Methods). Other dimensionality reduction techniques produce results that don't lend themselves to straightforward interpretations (such as PCA, which projects data to axes with no intrinsic meaning, or nonlinear dimensionality reduction), while other techniques cannot be applied to our experiment because of its nested design (as is the case of de-mixed PCA). Targeted dimensionality reduction uses the results of the linear regression to select individual directions in population space associated with any task variable of interest. In other words: of all the possible directions in population space, we choose the two most associated to covert attention or saccade direction, by picking the direction to which beta values from the linear regression, considered as a vector, point.

We computed directions maximally associated with covert attention and saccade direction but did so by a making minor but crucial modification to the original method. In order to test the priority map idea, which expects summation of different signals, we calculated the axes associated with cognitive variables using *only* conditions in which their associated stimuli are presented in the receptive field *alone*. That is, we used trials where only moving-dot surfaces were present in the receptive field to compute an axis associated with covert attention and, similarly, we used trials where only saccade targets were present in the RF to compute an axis associated with saccade direction. Using these axes, which should remain valid in a priority map, we plotted the mean population time courses on trials when *both* surfaces and saccade targets are present in the RF. All calculations were performed on neural activity aligned to saccade onset.

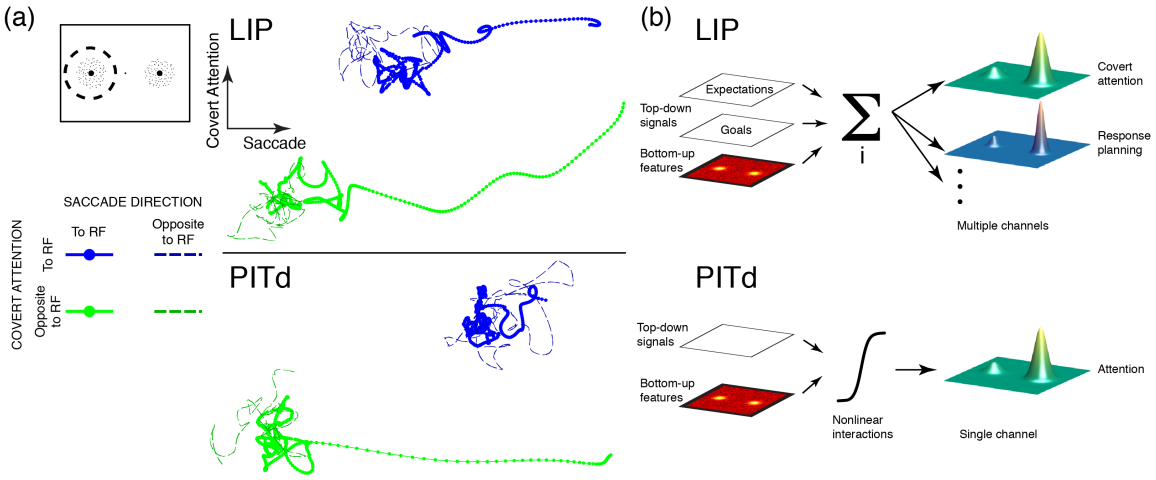


Figure 3.5

Population dynamics on task relevant axes and conceptual diagram.

(a) Mean population time courses from both monkeys combined in LIP (top) and PITd (bottom) projected on axes associated with covert attention condition (vertical) and saccade direction (horizontal), when both a dot surface and a saccade target are located inside the receptive field of cells. Color denotes covert attention condition and line style saccade direction (see inset). Units are z-scores and since they provide meaningful information are not shown. Crucially, units are the same in both axes, to ensure the shape of the traces is not altered, which is the important feature in these plots.

(b) Conceptual diagram illustrating coding of simultaneous variables in LIP through sum of disparate signals (top) and coding of attention in PITd (bottom), where nonlinear interactions dominate encoding.

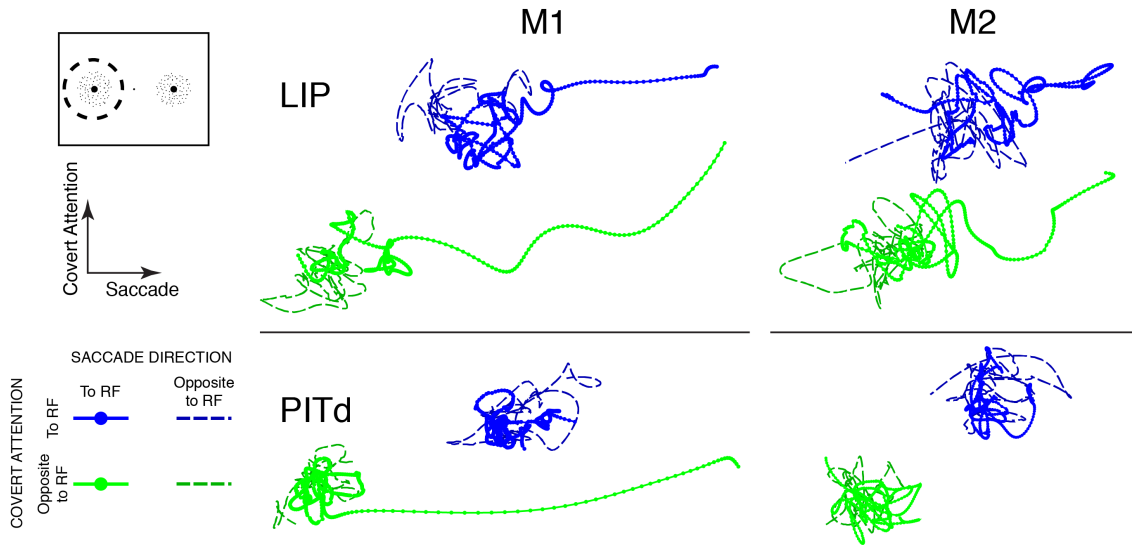


Figure S3.5

Population dynamics on task relevant axes in individual monkeys.

Mean population activity in monkeys M1 (left) and M2 (right) in LIP (top) and PITd (bottom) projected on axes associated with covert attention condition (vertical) and saccade direction (horizontal), when both a dot surface and a saccade target are located inside the receptive field of cells. Color denotes covert attention condition and line style saccade direction (see inset). Units (z-scores, not shown) are the same for vertical and horizontal axes.

Figures 3.5a and S3.5 (for individual monkeys) show the results of targeted dimensionality reduction in PITd and LIP. The figure shows the mean (i.e., across repeated trials) time evolution of the population for each covertly attended location (blue vs. red traces) and each saccade direction. The time courses start 1.5s and end 0s before saccade onset. During time periods when activity stays constant, the traces revolve around a fixed point due to small activity fluctuations. When they evolve and the state of the population changes, a characteristic trace of the trajectory is seen. In LIP there was a clear distinction between covert attention conditions along the axis associated with that variable, indicating its validity generalized from the trials with surfaces alone in the receptive field. In addition, the conditions in which saccade direction was directed towards the receptive field showed a buildup of activity along the saccade axis toward the end of the trial, confirming also the validity of the saccade axis computed from the condition with saccade targets alone in the receptive field. In PITd, covert attention conditions differed across the attention axis, showing that covert attention effects generalize across conditions. However, there was no reliable signature of saccade direction. In the combined results of both monkeys, and in monkey M1, a saccadic trace existed in the condition when covert attention is directed away from the receptive field. Yet, neither combined nor individual monkeys had an identifiable saccadic response when attention was directed toward the receptive field. This implies there is no generalization in PITd of results from saccades obtained with saccade targets alone in the receptive

field, only covert attention ones. In this brain area, covert attention and moving-dot surface presence effects dominate, and saccadic tuning is extinguished when combined with them, which manifests itself in the linear regression in the form of interactions.

These results provide a qualitative interpretation to the preceding ones using a linear regression and a decoding approach. In LIP, the results are fully compatible with a priority map and interactions, which exist, do not break down the code. This discards the alternative “pure intention” or “pure attention” scenarios. In the pure intention scenario, effects of covertly attended locations could have been observed when no saccade was being made. Crucially, right before the saccade, the presaccadic trace in both conditions would have been indistinguishable; since the motor plan is identical. Conversely, in a “pure attention” scenario, no distinct saccadic trace would have been observed in trials where covert attention is already at the saccade location. Instead, disparate signals are linearly combined (Figure 3.5b, top panel). This summing scheme has previously unrecognized consequences for the amount of simultaneous information contained, as revealed in the preceding decoding results. There was not one single priority signal, but rather multiple coexisting channels of information that have stable interpretations as being informative for covert attention and saccade planning. In PITd, attempts to identify separate channels failed (Figure 3.5b, bottom panel), because disparate signals nonlinearly interacted to the point that covert attention was robustly represented, at the

expense of other signals. This is consistent with a “pure attention” scenario, where covert attention is robustly signaled, and saccadic signals are observed only when the saccade is directed to a location not previously selected covertly. We discuss the imlications of these findings in Chapter 4.

Chapter 4:
Discussion and conclusions

We have presented experiments that illustrate two different aspects of visual processing. The first one, in Chapter 2, is an investigation of specializations for facial motion within the network of face-processing areas in the ventral stream. The second one, in Chapter 3, tests an influential hypothesis about the control of spatial attention in a ventral stream and a dorsal stream area.

Specialization for facial motion within the face-processing network

Our fMRI experiments in Chapter 2 show that dynamic stimulus information is processed differentially in the face-processing networks of two primate species. In macaques, face patches in the STS fundus were part of motion-selective cortex and exhibited enhanced responses to moving faces, which could be explained as a linear superposition of object preference and general motion sensitivity. This observation is consistent with two broad scenarios. It is possible that neurons within MF and AF respond to similar motion patterns as neurons outside MF and AF and yet differ from these in their preference for faces. Alternatively, MF and AF could consist of two different populations: one motion

selective, the other face-selective. The location and selectivity of MF raise the possibility that it might overlap with area LST (Nelissen et al., 2006), which exhibits both object and motion selectivity. However, as discussed below, a follow-up study by Fisher and Freiwald (2015) has rendered that scenario unlikely.

In contrast, face patches AL and AM responded selectively to facial motion. The interaction of shape and motion selectivity suggests that the two stimulus domains converge on the same population of neurons, perhaps receiving both face-selective and motion-selective inputs from AF and MF with whom they form a closed network (Moeller et al., 2008).

In humans, only the STS-FA showed a main effect of shape and motion, and a specific modulation by facial motion. This finding is consistent with results from Pitcher and colleagues (2011). In addition, only a contrast of moving faces versus moving objects reliably revealed the left hemisphere STS-FA, consistent with a study by Fox and colleagues (2009). These results support the idea that the human STS-FA is specialized for the processing of dynamic facial information, perhaps related in the left hemisphere to a sensitivity to lip movements caused by specializations for language (De Winter et al., 2015).

One motivation for these experiments was to use specialization for motion to shed light on putative homologies of face areas across species. Several interpretations of our results are possible; and, contrary to our expectation, they do not lend themselves to a straightforward equalization of face areas. The most

striking specialization is that of the human STS-FA for facial motion, unmatched by any of the other face areas, human or macaque. In macaques, specializations for facial motion are less prominent. It thus seems plausible that the human STS-FA might be a specialization of human or hominoid brains that other old world monkeys lack. The spatial separation of the STS-FA and its lack of connectivity with the other face areas (Gschwind et al., 2012) are compatible with this interpretation, but the positioning of macaque face areas inside or close to the STS has been suggested to imply the opposite assertion that macaque face areas might correspond to human STS-FA (Ku et al., 2011). However, considering the overall pattern of results in both humans and macaque monkeys of generally larger motion selectivity in more dorsal face areas (STS-FA in humans and fundus STS areas in macaque monkeys) than in more ventral areas, a homology of STS-FA with MF and AF is suggestive. This interpretation would be consistent with those drawn from processing of dynamic body shapes in the human and macaque brain that also found stronger motion selectivity in dorsomedial than ventro-lateral STS areas (Jastorff and Orban, 2009; Jastorff et al., 2012). As discussed, below a follow-up study (Fisher and Freiwald, 2015) found an additional face patch more anterior and dorsal than MF and AF, which could also be a homolog to the STS-FA.

In general, a word of caution is required regarding direct comparisons across species. One needs to consider several potentially complicating and limiting factors. First, although motion processing is important, multiple functional

dimensions should be considered for establishing homologies (Durand et al., 2009). Second, we focused our analysis on the face-patch system while dynamic stimuli are also processed outside of it (Furl et al., 2012; Jastorff et al., 2012; Nelissen et al., 2006). Third, contrast agents are typically used in macaques only. However, this particular difference is unlikely to obscure the comparison because MION and BOLD responses both ultimately tap into the same physiological mechanisms of neurovascular coupling. Fourth, the attentional state of macaques and humans is typically not fully controlled. We tried to limit this potential source of variability. The use of a simple task in humans encouraged them to distribute attention evenly across stimulation blocks, whereas in macaques extensive fixation training with the same stimuli presumably minimized fluctuations in internal state.

Our results underscore the importance of using naturalistic stimuli for studying functional areas. Dynamic faces elicited enhanced responses across all face-selective areas of the temporal lobe, and the left STS face area in humans was reliably active only for moving faces, whereas the functional differentiation we found in macaques face patches is, to our knowledge, the first one revealed with fMRI in this network. Crucially, their impact differs across face areas and thus helps to reveal functional differences within macaque and human face processing networks.

A follow-up study (Fisher and Freiwald, 2015) used dynamic stimuli to map face areas which led to the finding of a previously unreported face patch, MD,

that is anterior and dorsal to face patches ML and MF. Because of its selectivity to motion, MD could also be a homolog to the human STS-FA. While differences between their results and ours could in part be explained by potential differences in the motion energy of dynamic stimuli across studies, they found consistently with us, a division between more ventral STS face area and dorsal ones. In our experiment, that difference manifested itself in terms of the amount of motion responsiveness across more ventral and more dorsal face areas. They extended these results in a potentially insightful direction by showing selectivity for natural face motion as opposed to unnatural motion in the dorsal face areas. Together, the results suggest a distinction in the face processing-system that runs orthogonal to the hierarchical progression of increasing abstraction. More generally, motion appears to be an attribute than can be exploited for multiple goals, from the computation of changing spatial relationships in the dorsal visual stream, to the isolation of object structure or the understanding of socially relevant information in the ventral stream.

Spatial attention and priority in parietal and temporal cortex

Our experiments in Chapter 3 were designed to study a different question: how covert attention and eye movements are represented in areas that can control spatial attention.

Understanding how the brain encodes and utilizes cognitive spatial signals to guide perception and action is a major goal of cognitive neuroscience. At the behavioral level, it is clear that the brain can simultaneously accommodate signals necessary for covert attention of stimuli that need to be discriminated and signals for response selection that reflect commitment to a decision, even if they can spatially overlap with each other, but it is a mystery how this is achieved at the neural level. A number of brain areas, including LIP (Goldberg et al., 2012) and FEF (Armstrong et al., 2012) among others, have been proposed to represent both signals, but it is unclear how these two potentially conflicting messages can coexist in a given spatial map. Correlates of response selection and covert attention haven been studied in a number of areas but, in order to understand the specific contribution of each, it is necessary to study them under the same paradigms (e.g., Buschman and Miller, 2007; Hanks et al., 2015; Suzuki and Gottlieb, 2013). This issue becomes all the more pressing, as one of the biggest challenges the field faces is to explain how brain areas act as a network by supplementing each other or by division of labor to achieve a shared goal. An additional obstacle in this direction is that, until recently (Stemmann and Freiwald, 2015, in preparation), a systematic mapping of brain areas sensitive to spatial attention in monkeys was lacking, making it difficult to identify the key nodes of the network. This study addresses these questions by using a factorial design in which covert attention and response selection signals are studied either separately or in combination in brain areas identified by functional maps of the

macaque brain. We focused on two areas: one (PITd) for which there is only recent evidence that it can control at least certain aspects of visual attention (Stemmann and Freiwald, 2015, in preparation). Our experiment was designed to allow direct testing of an influential idea for framing spatial attention control, which proposed behaviorally relevant locations are highlighted by a general-purpose priority signal, which is computed through linear sum of disparate inputs (Bisley and Goldberg, 2010).

Our results suggest LIP, for which the idea was originally proposed to apply, uses the hypothesized encoding scheme. This, however, has previously unrealized consequences for the information present in the population, which can independently be decoded about diverse spatial cognitive signals. In this way, it is not necessary for downstream areas to determine when to apply a single priority signal to guide covert attention and when to use it to guide saccades, since this information is already explicit. The reason is a simple consequence of linear superposition: since each neuron weights each spatial signal differently, distinct channels of information emerge naturally as directions in population space.

A different scenario was revealed in PITd. First, we found tuned activity during the memory period of the MGS task, compatible with the existence of a priority signal. Attention is known to be sustained at the future saccade goal throughout the memory period of the MGS in both humans (Deubel and Schneider, 2003) and macaques (Bisley and Goldberg, 2003), and this might require an equally

sustained attention signal in the ventral stream. To our knowledge, this is the first report of such signals in the ventral stream where effects of attention in the have been previously reported to be visually gated (Ekstrom et al., 2008; Moore and Armstrong, 2003) or relatively modest baseline modulations (Luck et al., 1997). In our second task, however, we did not find summation of disparate signals, which suggests such scheme is not a general property of areas that encode spatial attention. PITd has been shown to satisfy most criteria to be understood as an attentional control area (Stemmann and Freiwald, 2015, in preparation). Yet, these signals are not independent. On the contrary, covert attention dominated in PITd, and could override any effect of saccade tuning, indicating covert attention is more fundamental in this area. The results suggest different levels of specialization across areas, with diverse spatial signals coexisting in LIP and robust signals for covert attention in PITd.

Efforts to separate covert attention from intentional responses often use antisaccade tasks where attention is presumably anticorrelated with eye movements or moves sequentially between locations. Our task design tried to overcome previous limitations, by requiring monkeys to make a difficult perceptual discrimination that demands attention and decision over extended periods of time. In this way, covert attention and response selection signals do not appear in sequence but rather simultaneously and independently. The results suggest the proposed encoding scheme to be valid in LIP, which rather than resulting in a unique priority signal gives place to two distinct simultaneous

spatial channels for covert attention and response selection. On the other hand, our PITd results indicate that in some cortical areas, covert attention for perception plays a more fundamental role than selection of the same location for action. In this context the existence of pre-saccadic signals can be expected to exist, as a consequence of the mandatory attention shifts that accompany a saccade. Importantly, the lack of distinct pre-saccadic tuning when an already covertly attended stimulus falls inside the receptive field of cells seems to confirm this role for PITd.

PITd, surrounded by feature-selective areas, is in an ideal position to transmit attention signals to surrounding feature-selective areas or to collect information about objects that is needed to compute priority elsewhere. The former possibility is compatible with a measurement of the timing of attention effects in ventral areas suggesting the effects of attention spread from some point close or possibly anterior to V4 (Buffalo et al., 2010; Mehta et al., 2000).

The present results are also of importance for studies of perceptual decision-making. Neural responses associated with response selection have been used in numerous studies as a window to the study of perceptual decision-making in several areas, chiefly including LIP (Gold and Shadlen, 2007). Claims about the neural correlates of decision-making based on these responses depend on them being robust to interference from other signals, particularly spatial ones. This is the first report to our knowledge, of distinct pre-saccadic signals to locations already selected by covert attention. Had responses in LIP

been similar to those in PITd in this respect, they would have casted serious doubts about the relevance of LIP responses in spatial decision-making. Our results suggest the possibility of a simple decoding mechanism that can take irrelevant modulations into account during readout of a decision-related signal.

Dimensionality of neural representations

Recently, Ganguli and colleagues (2008) proposed a model in which both covert attention and response selection signals in separate perceptual decision tasks evolve along a single dimensional manifold in neural space in LIP. Part of the allure of the proposal is that it showed how a single perceptual priority signal can be constructed from scaled spontaneous activity, leaving aside the question of whether disparate simultaneous signals can be represented. Our results suggest that covert attention and response selection vary along different dimensions such that both signals can coexist. This opens the question of whether there is an intrinsic dimensionality to the neural code, or it flexibly adapts to task demands. In our case, relevant dimensions could be extrapolated from task conditions in which a single relevant stimulus is present in the receptive field location, suggesting that responses under an increased load, such as when several task variables need to be represented by the same group of cells, exploit heterogeneity that is already observable in non-overlapping single tasks (Meister et al., 2013; Premereur et al., 2011).

On a related note, it has recently been suggested (Rigotti et al., 2013) that neurons in “higher” cortical areas benefit from producing high-dimensional representations that manifest themselves as non-linear interactions between task variables in the neural responses (non-linear mixed selectivity). This would contribute to behavioral flexibility by allowing information about any combination of task variables to be explicitly available, which could allow learning of practically unlimited new associations. In our task, we observe that linear rather than non-linear mixed selectivity dominates in parietal cortex, similar to other studies (Raposo et al., 2014; Rishel et al., 2013) investigating simultaneous representation of spatial and non-spatial signals. However, distinguishing between some task elements, like modality, was not required by the behavioral tasks (Raposo et al., 2014) and causal interventions found parietal cortex not to be necessary for non-spatial aspects of performance (Balan and Gottlieb, 2009) Because of that, the simultaneous representation of spatial signals in LIP, for which the area is necessary, remained an open question. Our finding of linear mixing of spatial signals in LIP suggest a different underlying logic between parietal cortex and prefrontal areas (Mante et al., 2013; Rigotti et al., 2013). A reason for this difference might be that the computational role of LIP is not to provide an exhaustive vocabulary of neural responses from which every desirable association can be learned, but rather to provide information about different aspects of behavioral priority in a format that allows straightforward generalization and robust readout. In this view, flexible responses from prefrontal

cortex could map to parietal cortex in a format that is already “digested” and ready to be used to influence sensory processing and motor responses. Downstream areas might not need to have available to them arbitrary combinations of task variables (such as a XOR function), only certain generally important variables. It would be certainly interesting to test responses in some version of our paradigm in prefrontal areas and study the construction of priority maps from abstract, flexible mixed selectivity.

Two streams of visual processing and distributed computations

Traditionally, the primate visual system has been divided into a ventral stream (where PITd belongs) and a dorsal one (where LIP is located) (Ungerleider and Mishkin, 1982), associated with vision for action and vision for perception respectively (Goodale and Milner, 1992). In recent years, the functional distinction between them began to blur with the report of object selective responses in the dorsal stream (e.g., Konen and Kastner, 2008; Sereno and Maunsell, 1998). Here, we report robust encoding of spatial attention signals in PITd, including sustained selectivity in a memory-guided saccade task, up to now a hallmark of the dorsal stream (Gnadt and Andersen, 1988), prefrontal areas (Bruce and Goldberg, 1985; Funahashi et al., 1989) and subcortical oculomotor structures (Glimcher and Sparks, 1992), suggesting involvement of ventral stream areas in coding spatial cognitive signals. We do not believe, however, the

distinction of ventral and dorsal streams to be affected by it. Considered as supporting vision for perception and for action, the responses reported in both streams are illuminated by that distinction, with responses in parietal cortex being greater to objects that can be grasped (Mruczek et al., 2013) and, conversely, spatial attention signals providing a robust representation of stimuli that need to be discriminated rather than for action.

Perhaps the best way to frame the present results in is in terms of a network of distributed computations, which are highlighted by simple experiments like the memory-guided saccade task. It is known, for instance, the responses during memory-guided saccades in LIP and FEF are interdependent (Chafee and Goldman-Rakic, 2000), making it clear that areas do not work in isolation but as a coherent whole. Conversely, it has been shown that cortical attention-like modulations can exist separately from the ability to select an object for behavior (Zénon and Krauzlis, 2012). A future challenge is to dissect these signals further and see which ones, like response selection activity in LIP or covert attention signals in PITd, survive manipulations in other areas. Only forcing the cognitive apparatus with a high load task, causal manipulations and parallel recordings a picture of mutual interactions during distributed computations will be revealed.

Materials and Methods

fMRI experiments

All animal procedures complied with the National Institutes of Health *Guide for Care and Use of Laboratory Animals*, regulations for the welfare of experimental animals issued by the California Institute of Technology, where all fMRI macaque experiments were conducted. All human subject procedures were approved by the Institutional Review Board of The Rockefeller University, and informed consent was obtained from all human subjects.

Surgery.

Implantation of MR-compatible headpost (Ultem; General Electric Plastics), MR-compatible ceramic screws (Thomas Recording), and acrylic cement (Grip Cement, Caulk; Dentsply International) followed standard anesthetic, aseptic, and postoperative treatment protocols.

Monkey fMRI

Scanning was performed on a 3T MR scanner (TIM Trio with AC88 gradient insert; Siemens). For each monkey, we acquired 16 anatomical volumes at high spatial resolution (0.5 mm isometric) with a T1-weighted inversion recovery sequence (MPRAGE) under anesthesia (ketamine and medetomidine, 8 mg/kg and 0.04 mg/kg). For functional imaging, contrast agent ferumoxytol (8 mg of Fe per kg body weight), was injected into the femoral vein before the scan session to increase the signal-to-noise ratio. Like MION (Vanduffel et al., 2001), ferumoxytol reduces signals in activated voxels, and we thus inverted signals for display of functional data to facilitate comparison with BOLD data.

All functional data were acquired in horizontal slices with a multiecho EPI sequence (TR 2 or 3s, TE 30ms, 1.5 or 1.0 mm3 voxel size) and a custom-made 1-channel or 8-channel surface coil as described previously (Tsao et al., 2008). The use of smaller voxel sizes in macaques reduces the effect of ear-canal-related susceptibility artifacts compared with humans (Devlin et al., 2000; Kriegeskorte et al., 2007). Three male rhesus monkeys (*Macaca mulatta*) were scanned while foveating a fixation dot at the center of the screen. Monkeys sat in sphinx position with their heads fixed (Tsao et al., 2003; Vanduffel et al., 2001). Juice reward was delivered after variable periods of time (2–4 s) during which the monkeys maintained fixation within 2 degrees of the fixation dot. Eye position was measured at 100 Hz using a commercial eye monitoring system (ISCAN).

Human fMRI

All scanning was performed on a 3T MR scanner (TIM Trio; Siemens). Human functional data were acquired in horizontal slices, approximately aligned to the AC-PC line with a standard EPI sequence (TR 2 s, TE 32ms, 64 x 64 matrix, 3.43 mm x 3.43 mm in-plane resolution, 3.4 mm slice thickness, flip angle 90°) and a 32-channel head coil. On each scan session, we obtained a high-resolution anatomical volume of the entire brain (MPRAGE, 1 mm isometric).

Six human subjects (3 females, 3 males; age 25–35 years) participated in the experiment. Subjects were instructed to maintain fixation on a central dot and indicate with a button press (right index finger) when the identity of a stimulus was repeated within a visual stimulation block. Eye position was measured at 100 Hz using a commercial system (ISCAN) to ensure that subjects were following fixation instructions within a 2-degree window.

Visual stimulation

The same visual stimuli were presented to humans and macaque monkeys. Two different experiments were performed, both presented in block designs in separate runs.

The first experiment was a standard face localizer (Moeller et al., 2008), used to define face-selective ROIs. The duration of each block was set to equal 8 times the TR of the imaging sequence. Each image block contained pictures of one of the following categories: human faces (F), monkey faces (M), human hands (H),

gadgets (G), fruits and vegetables (V), and human headless bodies (B). Each image block was preceded by a scramble block (S) with spatially scrambled versions of the pictures of the subsequent block. Runs concluded with a final block showing a gray random dot pattern (R). Thus, the sequence of blocks presented in each run was S F S H S M S G S F S V S M S B R. Images were presented at a subtended 5.9° visual angle (10.4 cm diameter at 100 cm distance) for 0.5 s.

The second experiment was performed to test for selectivity for stimulus dynamics. Blocks lasted 32 and 30s for all scans. Stimulus conditions were comprised of blocks of moving faces, static faces, moving objects, and static objects. Motion blocks were composed of short movies (0.5–2.5 s long), whereas static blocks included pictures shown for the same amount of time. Face movies showed macaques or humans vocalizing and generating facial expressions. Macaque facial expressions included coo calls, lip smacking, aggressive teeth displays, and grunts. Human expressions included smiling, nodding, and simple vocalizations (similar to monkey calls). Object movies showed artificial and natural objects (computer mouse, shoe, canned food, tooth- brush, comb, flowers, leaves, fruits) subject to naturalistic motions (such as falling or being shaken as if moved by the wind or sliding down a slope).

To minimize low-level differences across stimuli, images and movies were achromatic, objects and faces were placed in the picture center and on identical backgrounds of salt-and-pepper noise, and movies and pictures were manually

adjusted to have an overall matched distribution of pixel intensities and similar object/face sizes. Motion energy is hard to measure in naturalistic movies. We compared activations in general motion-sensitive brain areas (see below) to estimate differences across motion blocks. Static control conditions were generated directly from the corresponding movie clip by extracting frames maximizing the social information conveyed. For instance, if the original clip showed a monkey with an aggressive expression, we used a frame with the teeth most visible. Image categories comprised static human faces (FHS), static objects (two sets, OS and OSbis), moving human faces (FHM), moving objects (two sets, OM and OMbis), static monkey faces (FMS), moving monkey faces (FMM), and scrambled images (S). The sequence of blocks used was as follows: S FHS S OS S FHM S OM S FMS S OSbis S_FMM S OMbis S. Stimuli subtended 7.4° visual angle (13 cm diameter at 100 cm distance).

Visual stimulation was controlled by custom MATLAB (MathWorks) code using the Psychophysics Toolbox (Brainard, 1997). Stimuli were projected with a video projector (JVC DLA-G15E) at 30 Hz with 720 x 480 pixel resolution on a back-projection screen.

fMRI data analysis

FreeSurfer and FSFAST (http://surfer.nmr.mgh.harvard.edu/) were used to reconstruct cortical surfaces and perform functional data analysis, following procedures detailed previously (Tsao et al., 2003). The same procedure was

used to define face-selective areas in monkeys and humans. We used data from Experiment 1 and calculated the contrast of static faces versus all whole objects. Face-selective regions were identified by anatomical location and relative position. Identity of face-selective regions was then determined by comparison with published coordinates (Pitcher et al., 2011; Tsao et al., 2008), and established naming conventions were used. Intensities from voxels within ROIs were pooled together for subsequent analysis of data from Experiment 2. We used high thresholds (at least $p < 10^{-7}$) to define macaque ROIs to minimize partial volume effects. An exception is left middle face patch in the STS fundus (MF) in Monkey M1 ($p < 10^{-2}$) where selectivity was confirmed in Experiment 2.

In macaques the number of runs used depended on individual performance and varied somewhat: 22, 21, and 28 runs in Experiment 1 and 28, 21, and 16 in Experiment 2 for Monkeys M1, M2, and M3, respectively. In humans, we had 4 runs per subject and experiment.

For group analysis, a general linear model was fit to the β values obtained from every single run of Experiment 2 for each ROI. Because of the small sample size (3 subjects), typical for monkey fMRI studies, a fixed effects group analysis was used as in previous studies (Jastorff et al., 2012). Contrasts were computed with a two-way ANOVA, with single-run β values as repeated measures. Bonferroni corrections for multiple ROIs were used to adjust significance thresholds.

Motion energy controls

Possible differences in motion energy between face and non-face object stimuli were assessed via activation differences in motion-sensitive areas in the STS of macaques. We identified MT/MST/FST by contrasting all moving stimuli with all static ones on even runs, setting a high significance threshold, and verified results by registration to macaque F99 atlas (Van Essen et al., 2011). For these ROIs, we contrasted moving versus static faces and moving versus static objects on odd runs to measure modulation by face motion (7.02 ± 0.43%, mean ± SEM; $F(1,429) = 272.34$, $p < 10^{-47}$) and by object motion (9.76 ± 0.60%; $F(1,429) = 263.19$, $p < 10^{-45}$). This suggests that motion energy in non-face movies was higher than in face movies. Thus, larger activations in a face area for facial versus object motion cannot be explained by differences in motion energy.

Eye movement controls

To assess whether eye movements were different across conditions, we calculated the number of saccades during the task. Eye traces were low-pass-filtered (15 Hz cutoff frequency) and underwent edge-preserving smoothing for noise removal (Santella and DeCarlo, 2004), after which a velocity threshold was applied. Blinks were excluded from the analysis. For each monkey, we performed a one-way ANOVA on the number of saccades during each stimulation block using different runs as repeated measures. Only Monkey M2 showed a significant effect of condition on saccade number ($F(3,132) = 4.36$, $p < 0.01$),

whereas Monkeys M1 ($F(3,172) = 2.57$, $p = 0.06$) and M3 ($F(3,116) = 1.51$, $p = 0.22$) did not. A *post hoc* analysis in Monkey M2 showed no significant difference within motion conditions.

Electrophysiology of attention

All animal procedures complied with the US National Institutes of Health *Guide for Care and Use of Laboratory Animals* and were approved by The Rockefeller University Institutional Animal Care and Use Committee (IACUC).

Subjects and Surgical Procedures

Experiments were performed with two adult male rhesus macaques (*Maccaca mulatta*), weighing 8-10 kg. Implantation of MR-compatible headpost (Ultem; General Electric Plastics), MR-compatible ceramic screws (Thomas Recording), and acrylic cement (Grip Cement, Caulk; Dentsply International) followed standard anesthetic, aseptic, and postoperative treatment protocols.

MRI-guided electrophysiology

Electrophysiological recordings in both animals were guided by statistical maps of the effects of covert spatial attention obtained in a previous study (Stemmann and Freiwald, 2015, in preparation). In brief, animals were trained to perform an

attention-demanding covert motion discrimination task, related to the one used in the present study. They were required to covertly attend one out of two moving-dot surfaces (MDSs), positioned in the left and right hemifields along the horizontal meridian. Animals reported the direction (eight possibilities) of a prolonged motion event in the cued surface, with an eye movement to one out of eight possible saccade targets. A contrast of activity on attend-left vs. attend-right conditions was used to create a statistical map which served to estimate the location of attention-sensitive hotspots in areas PITd and LIP to be targeted with recording electrodes (Figure S3.1). Vertical Plastic recording chambers were positioned vertically to reach PITd and LIP. In monkey M2, a second lateral chamber was used in PITd recordings to avoid blood vessels. We used the Planner software (Ohayon and Tsao, 2012) to calculate the angle and position of desired electrode trajectories from MR anatomical highlighting showing blood vessels to be avoided and functional maps. We used plastic grids, available commercially (Crist Instruments) or 3D-printed from plans created by the Planner software placed inside plastic recording chambers.

Extracellular recordings

Extracellular recordings were conducted using single Tungsten electrodes (FHC, impedance 2-9 MΩ/1kHz). Electrodes were back-loaded into metal guide tubes of length set to reach, from the top of the grid holes, approximately 2mm below the dura and were slowly advanced using a manual oil hydraulic manipulator

(Narishige Scientific Instruments). The electrophysiological signal was amplified and waveforms that crossed a set threshold were sorted online into separate units using multiple discrimination windows or ellipsoids defining clusters in principal component space. Spiking activity, local field potentials, eye position traces and digital triggers of task events were recorded using a Cerebus data-acquisition system (Blackrock Microsystems). Spike isolation was assessed offline by existence of an absolute refractory period in inter-spike interval histograms, degree of overlap in principal component space and inspection of waveforms. Each well-defined cluster was treated as a 'unit' for the purposes of the analyses.

Behavioral monitoring and stimulus presentation

Behavior was controlled and stimuli were presented using custom software 'Visiko' written in C++, running on a Windows PC that received and sent signals via an analog and digital input/output card PCI-DAS1002 (Measurement Computing Corporation). The software controlled juice rewards and sent triggers to the Cerebus data-acquisition system. Eye position was measured and recorded at 100Hz using an infrared eye-tracking system (ETL-200, ISCAN Inc.). Stimuli appeared on a CRT (cathode ray tube) computer monitor (36.6 x 27.4 cm; 1920 x1440 pixels; 100 Hz refresh rate) at a distance of 57 cm from the eyes.

Cell selection

During recording sessions, the presence of neural activity was monitored by the sound of the amplified neural signal connected to an audio monitor and by observing the presence of waveforms on a computer screen. The consistency of transitions between gray matter, white matter and sulci as assessed from these measures was compared with expected trajectories in anatomical MR images, while expected visual, auditory or somatosensory responsiveness of areas along electrode trajectory was verified and compared with that expected from brain atlases and preceding recording sessions. In LIP recordings, activity during memory-guided saccades was used to verify that the final electrode position was correct during the first recordings in a given recording site, but this activity was not a requirement for a given cell to be recorded. In PITd recordings, the electrode was advanced until non-foveal visual responses were observed. We did not select cells based on any requirement other than visual tuning. On a typical session we would first record responses during the receptive-field (RF) mapping task, used to tune stimulus parameters (positions) in subsequent tasks. Then we recorded responses during a memory-guided saccade (MGS) task and finally the spatially cued discrimination task (described below).

Receptive-field mapping task

We used this task to compute a quantitative spatial map of a cell's RF. Monkeys were required to fixate on a white central spot (diameter: 0.5 degrees of visual

angle (º), fixation tolerance window: 4deg, luminance: 26cd/m^2, CIE chromaticity: x=0272 y=0.274) on a dark gray background (luminance: 2.53cd/m^2, CIE chromaticity: x=0.251, y=0.217). Reward was delivered after a variable period (3-3.5s) of sustained fixation. White squares (100ms duration separated by a 300ms pause, width: 2deg, luminance and chromaticity as fixation spot) were presented at random positions along an equilateral triangular lattice (spacing: 1.5deg, extent: 30deg) covering the computer screen. Experiments typically lasted 5-7min.

Memory-guided saccade task

Monkeys had to make a saccade to a remembered location in the absence of a visual stimulus (Figure 3.1d). Subjects initiated a trial by fixating on a central blue dot (fixation tolerance window: 5deg, diameter: 0.5deg, luminance: 6.3cd/m^2, CIE chromaticity: x=0115 y=0.116) on a black background (luminance: 0.01cd/m^2, CIE chromaticity: x=0.125, y=0.124). A central fixation period (300 to 600ms) was followed by a peripheral visual target presentation (duration: 200ms, diameter: 0.8deg, saccade window: 7-10deg, luminance: 26cd/m^2, CIE chromaticity: x=0.272 y=0.274) Target locations had a fixed eccentricity (selected to match that of the RF of the cell being recorded) and one of eight equally separated polar angles in visual space (meridians and diagonals). After extinction of the target, monkeys waited for a variable period of time (memory period, 900-1400ms) until the central fixation spot disappeared. Monkeys were rewarded with a juice drop if

they made a saccade to the remembered target (~7deg variable tolerance window) location within 400ms of disappearance of the central fixation spot. After a correct response, the saccade target briefly reappeared (100ms). Typical experiments included 100 correct trials (roughly 12 repetitions per target location). Measurements of the dynamics of luminance of the saccade target with a photodiode confirmed it lasted 200ms, after which luminance decayed to baseline levels in a time shorter than the monitor inter-frame interval (10ms).

Spatially cued discrimination task

Monkeys had to detect a prolonged motion event (PME) in one out of two moving-dot surfaces and report its direction with a saccade. Trial structure is schematized in Figure 3.1e. Trials started with a 1s central fixation period (diameter: 0.25deg, fixation tolerance window: 4deg, luminance: 2.5cd/m^2, CIE chromaticity: x=0.251 y=0.217), on a black background (luminance: 0.01cd/m^2, CIE chromaticity: x=0.125, y=0.124). Saccade targets were present throughout the trial (annuli inner diameter: 0.2deg, outer diameter: 0.3deg, saccade window: 5-8deg, luminance: 2.5cd/m^2, CIE chromaticity: x=0251 y=0.217). A small central rectangular cue (length: 0.25deg, width: 0.07deg, luminance and CIE chromaticity as fixation spot) appeared pointing in the direction of the future relevant moving-dot surface, staying during the whole duration of the trial. After 1s, the two moving-dot surfaces (diameter: 4-5deg, mean luminance: 0.12cd/m^2. Dot size: 0.2deg, density: 6dots/deg^2, speed: 6deg/s, lifetime: 100ms, luminance:

1.2cd/m^2, CIE chromaticity: x=0.245 y=0.191) appeared on opposite sides of the visual field, the cued surface being behaviorally relevant while the other one was a distractor that had to be ignored. A fixed fraction of dots moved in a consistent direction (motion coherence 9-28%, adjusted in each session). For each surface, motion direction changed randomly and independently (24 possible directions) every 60ms for a variable number of times (1 to 60), independent for each surface. These irrelevant translations were followed by the PME during which the direction of motion remained the same (2.7s). During the PME, dots moved into one of two opposite directions. Monkeys had to indicate the direction of motion of the cued surface's PME with an eye movement to the corresponding saccade target. As mentioned above, motion coherence levels varied from session to session to keep performance levels around 66-75% correct. In this way we compensated for performance variability, which tended be affected by the use of different stimulus eccentricities.

The goal of the task design is to study spatial signals related to covert attention to the moving-dot surfaces and response selection with a saccade, when both are spatially separated or allowed to overlap. To do that, we had conditions in which there was only a dot-surface in the RF of the recorded cell, or only a saccade target in the RF, or both, or none (Figure 3.1e, right panel). Each of these four stimulus arrangements included four behavioral conditions (two possible cue directions and two possible saccade directions), thus generating a total of 16 unique conditions overall. The two surfaces were always positioned on

opposite sides around the central fixation point and the same was true for the two saccade targets. Depending on the trial condition, moving-dot surfaces could be aligned with saccade targets or lie along a line perpendicular to them. Stimulus eccentricity was arranged to match that of the RF, polar angles were chosen between meridian and diagonal, whichever matched the RF polar angle best.

The different conditions were presented in blocks of trials that had the same stimulus arrangement and cued surface, but randomized PME timing and direction. A block was completed when six trials had been performed correctly.

Several aspects of task design ensured that monkeys needed to gather relevant information from the cued surface as opposed to randomly guessing, getting useful information from the distractor surface, or responding in an impulsive or stereotyped way. First, in order for responses to be considered correct monkeys had to make a saccade only after the PME on the cued surface started, which reduced chance performance levels well below 50% if an eye movement was made at random times. Second, the onset time of the PME varied within a long time window, which made it difficult for monkeys to randomly guess when a response was appropriate. Third, 10% of trials were “catch” trials that occurred randomly within the blocks, where no prolonged motion event occurred in the cued surface, and monkeys were rewarded for maintaining fixation until the end of the trial. Fourth, the PMEs in the cued and distractor surfaces were uncorrelated both in time and motion direction, ensuring no useful information about either timing or motion direction could be extracted from the distractor

surface. Fifth, during training sessions preceding recordings only, incorrect trials were repeated until the monkey provided a correct response so as to make stereotyped responses an inviable strategy. Sixth, trials with incorrect responses were punished with an increased inter-trial time. Seventh, reward magnitude was incremental, i.e. increased after a successful trial and reset to initial levels after a mistake, encouraging monkeys to correctly perform many trials in a row.

Analysis of neural responses: Receptive Field Mapping

For each stimulus position, the mean firing rate across repeated presentations (including only those in which there was successful central fixation) was calculated using a temporal window from 50 to 150ms after stimulus onset. Firing rates were interpolated to all positions in visual space via radial basis method. Results were smoothed via convolution with a Gaussian kernel (2deg FWHM, truncated at 3deg). The RF center was estimated, after z-scoring the resulting map, removing points with $z<0$, and determining its largest connected component, by calculating its center of mass.

Analysis of neural responses: Memory-guided saccades

All analyses were performed in coordinates relative to the RF, either estimated from the RF mapping task or, in rare exceptions when the acquired map was deemed too noisy for an accurate estimate of the RF center, from the target position that elicited the largest visual response. For each trial, we calculated a

spike-density function (SDF) by convolving each spike (considered as a delta-function) with a Gaussian kernel (σ=25ms, truncated at 100ms). Activity was aligned to two events: target onset and saccade onset. Target-onset-aligned activity included time from 500ms before target onset until 500ms before saccade onset. Saccade onsets were defined as the time point the eye position left the fixation window (2 degree wide). Saccade-onset-aligned activity included spikes from 400ms after target onset until 200ms after saccade onset. Because trials have different durations, a variable number of trials contribute to the mean time course in a given task condition. For quantifying the strength of effects across the population, the activity of each cell was z-scored as follows. First, for each of the 8 target positions, the mean time course across repeated trials was computed, to ensure all conditions were equally represented independently of the number of trials available. Second, the mean activity and standard deviation across all 8 target positions and time points was computed. Third, the mean was subtracted from the SDF, which was then divided by the standard deviation. To quantify tuning to target position, a linear regression was performed to trials in which the target was positioned inside the RF and trials in which it was opposite to it, using target presence in the RF as the only predictor. This can be done for each time point or using the average in a given time window. To compare the distribution of tuning strength during the memory period, we used a time window from -800ms to -600ms relative to saccade onset. To quantify the existence of a pre-saccadic increase of activity we computed the difference of activity in the previous time

window vs. one from -200ms to 0ms relative to saccade onset, using only trials were the target had been presented in the RF. We used two-tailed t-tests to quantify the significance of results.

Analysis of neural responses: spatially cued discrimination task - General procedures

As above, all analyses were done in coordinates relative to the RF center. SDFs were calculated as in the MGS task. Activity was aligned to two events: moving-dot surface onset and saccade onset, the latter corresponding to the time point where gaze leaves the fixation window. The surface-onset-aligned analysis window ranged from 1s before moving-dot surface onset to 500ms before saccade onset. Saccade-onset-aligned activity included spikes from 300ms after moving-dot surface onset until 600ms after saccade onset. As in the MGS task analyses, activity of each unit was z-scored for estimating effects across the population: average time courses were calculated for each of the unique 16 conditions, global mean and standard deviation were calculated across all time points and conditions. Then, SDF in each trial and time point had the global mean subtracted and was divided by the global standard deviation.

Analysis of neural responses: spatially cued discrimination task - Linear regression

To quantify the effect of each task variable on the responses of each recorded unit, we used a multi-variable, linear regression model (Mante et al., 2013; Rorie et al., 2010). The model finds, for each time point, the linear combination of experimental variables that best explains the observed responses in a given neuron, in the sense of minimizing the sum squared errors. In its simplest version, which we show here for conceptual simplicity, activity is described as

$$r_{i,t}(k) = \beta_{i,t}(0) + \beta_{i,t}(1)\, surface(k) + \beta_{i,t}(2)\, target(k) + \beta_{i,t}(2)\, attention(k) + \beta_{i,t}(2)\, saccade(k) \quad (1)$$

where $r_{i,t}(k)$ is the firing rate (z-scored) of neuron *i* at time *t* in trial *k*, $surface(k)$ indicates the presence of a moving-dot surface in the RF (+1/2: present, -1/2: absent), $target(k)$ refers to the presence of a saccade target in the RF (+1/2: present, -1/2: absent), $attention(k)$ denotes if covert attention is directed to the RF (+1/2: to RF, -1/2: opposite to RF, 0: surface not present in RF), $saccade(k)$ specifies saccade direction (+1/2: to RF, -1/2: opposite to RF, 0: saccade target not present in RF).

The regression coefficients $\beta_{i,t}(v)$, for $v = 1, \ldots, 4$ quantify how much of the z-scored firing rate depends on a given task variable ($v = 1$: surface presence, $v = 2$: target presence, $v = 3$: covert attention, $v = 4$: saccade direction). The first

regression coefficient ($\nu = 0$) reflects the time-dependent component of the firing rate that is not dependent on any of the task variables. Equation (1) reflects the simplest possible linear regression of the firing rate, but we also included similar terms (not shown here for simplicity) that captured the effect of covert attention to different positions when the moving-dot surface was not inside the RF and an analogous term for saccade direction when saccade targets were not positioned in the RF. In addition, while additional terms for interactions between experimental variables can be added, we chose to calculate the size of interactions as a "difference of differences". For instance, the interaction of saccade direction and covert attention was the difference between the effects of saccade direction when attention was directed to the RF vs. when attention was opposite to it. To compute the different interactions in this way, corresponding reparametrizations of the linear regression were used. Similarly, since main effects are hard to interpret in the presence of interactions, our results section reports the single effect of spatial attention, computed in trials where saccade targets were not in the RF, as opposed to the main effects from equation (1). Analogously, we report the effects of saccade direction for trials in which moving-dot surfaces were not present in the RF, rather than the main effect of saccade direction as computed in equation (1).

The linear regression was performed independently at each time point to visualize the dynamics of the different effects. For quantification of effects across cells, we performed the regression using fixed time windows: 200 to 0ms before

saccade onset for saccade tuning and its interactions and 800 to 500ms for all other effects and associated interactions. All statistical tests were two-tailed t-tests, computed using Free Surfer's FSFAST toolbox for MATLAB (http://surfer.nmr.mgh.harvard.edu/).

To calculate the time at which selectivity for saccade direction started to emerge, we considered all cells that showed significant saccade tuning (two-tailed test, $p<0.05$) in the last time point before the saccade. Trials in which saccade targets were the only stimulus in the RF were used. For each time point (3.3ms sampling period), we computed the significance of saccadic tuning. Onset of saccadic tuning was defined as the first time point, counting in reverse order from saccade onset, where saccade tuning was significant. As in other studies (Ipata et al., 2006), we controlled for random fluctuations by restricting the analysis only to cells where saccade tuning was significant in several consecutive time points. Only cells in which the effect of saccade direction remained significant for at least 100ms (30 time samples) were considered in this analysis.

Analysis of neural responses: spatially cued discrimination task - Decoding analyses

We quantified the presence of distributed information in the population about covert attention and saccades using a decoding approach, in which we trained a

decoder using the activity of pseudo-populations of cells in each area. Decoding analyses were performed using the Neural Decoding Toolbox (Meyers, 2013), a MATLAB package implementing neural population decoding methods. Activity of all units was considered relative to their receptive field, so as to form a collection of units that encode information about the same portion of space, similar to what is expected from cells belonging to the same cortical column. We performed these analyses on data aligned to saccade onset, binned every 100ms. We restricted decoding to units for which at least 6 repetitions from each task condition used in each analysis was available. Each subpopulation was composed of the same number of units, randomly sampled with replacement from the pool of cells available from each area.

We followed a cross-validation approach, in which a subset of the data (83% of available trials in each condition) was used to train the decoder, while the remaining trials (17%) were used to test the decoding accuracy, as a measure of the reliability of the learned relationship between neural activity and experimental condition. For each unit, data were randomly selected from 6 trials from each of the relevant conditions. For each of these trials, data from 20 units was concatenated to create pseudo-population response vectors (that is, data from units recorded mostly on separate sessions but treated as if they had been recorded simultaneously). The pseudo-population vectors were grouped into 6 splits of the data, each containing a response vector for each relevant condition. A support vector machine (Cortes and Vapnik, 1995) was trained using 5 splits of

the data and the performance of the classifier was tested using the remaining split of the data. Before providing the data to the classifier, data from each unit was z-scored (using parameters estimated from the training set), to prevent neurons with high firing rates from dominating the representation of information. The same procedure is repeated 6 times, each time selecting a different split of the data as test set. In other words, a 6-fold leave-one-split-out cross-validation procedure was used. As a measure of performance, we calculated the fraction of test examples that were correctly classified (i.e., zero-one loss measure). To get a more precise estimate of the decoding performance as well as an estimation of the variability of the results that would arise as a consequence of selecting different subsets of neurons drawn from the same original population, we repeated the decoding procedure 200 times, using a bootstrap method. On each repetition, cells were drawn randomly from the available pool of cells with the possibility of drawing the same cell more than once (random sampling with replacement). The results we report are averaged across the 200 repetitions.

To determine statistical significance of results, we used permutation tests. Decoding procedures were repeated as before, but labels of experimental conditions were randomly permuted before supplying them to the decoder. We used 50 rather than 200 repetitions for each permutation to reduce computation time. This makes the test more conservative since it overestimates the variance of the distribution, relative to the use of 200 repetitions. As a consequence, the probability of a type II error was increased, but not that of a type I error.

Analysis of neural responses: spatially cued discrimination task - Extraction of task-relevant dimensions

To visualize the dynamics of the populations along dimensions most informative about task variables, we used a recently developed dimensionality reduction technique, targeted dimensionality reduction (Mante et al., 2013). The average SDFs from all cells in a given area are concatenated to form a population SDF vector. SDF vectors are “de-noised” by performing a PCA and reducing the dimension of the vector to a smaller number of informative dimensions. In our case we used 12 PCs, following Mante and colleagues (2013). Targeted dimensionality reduction isolates a dimension associated with any task variable of interest by considering regression coefficients like those of equation 1 (see above) of a collection of units as a vector. The regression coefficients of all units for the task variable of interest are concatenated forming a vector, the dimension of which equals the number of units. This vector is then projected to the same PC space computed above from SDFs. Since regression coefficients evolve in time, so does their associated vector. The direction associated with a task variable of interest is the direction of its regression coefficient vector in PC space at the time point where the norm of the vector reaches its maximum value. It should be noted that when more than one variable of interest is isolated, the associated directions are in general not orthogonal. This is problematic because the projection of activity to the subspace spanned by these directions, although

unique, is not simply the independent projection to each of the computed directions. This can be solved by an orthogonalization procedure (such as a Gram–Schmidt process), which leaves the spanned subspace unchanged. The multidimensional SDFs can then be linearly projected into the resulting axes for visualization.

We made a minor but important modification to the targeted dimensionality reduction technique. The priority map hypothesis proposes that neural activity in the presence of multiple task factors is roughly the sum of responses to each task factor in isolation (Ipata et al., 2009). To test this hypothesis, we computed directions associated with covert attention and to saccade direction from the *single* effects and then used these directions project the population SDFs during the simultaneous presence of moving-dot surfaces and saccade targets in the RF. If linear summation of disparate sources occurs, then direction computed from single effects should remain valid when plotting SDFs from a different condition. This modification has the dual advantage of testing the proposed encoding scheme for priority maps and, in addition, serving as a cross-validation procedure ensuring the selected directions do not reflect the particular realization of the noise in the conditions plotted. All calculations were done using neural activity temporally aligned to saccade onset.

References

Allison, T., Puce, A., and Mccarthy, G. (2000). Social perception from visual cues: role of the STS region. Trends Cogn Sci (Regul Ed) *4*, 267–278.

Andersen, R.A., Asanuma, C., Essick, G., and Siegel, R.M. (1990). Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. J Comp Neurol *296*, 65–113.

Andersen, R.A., Snyder, L.H., Bradley, D.C., and Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Annu Rev Neurosci *20*, 303–330.

Andersen, R.A., and Buneo, C.A. (2002). Intentional maps in posterior parietal cortex. Annu Rev Neurosci *25*, 189–220.

Armstrong, K.M., Schafer, R.J., Chang, M.H., and Moore, T. (2012). Attention and action in the frontal eye field. In The Neuroscience of Attention, G.R. Mangun, ed. (Oxford University Press New York), pp. 151–166.

Asaad, W.F., Rainer, G., and Miller, E.K. (1998). Neural activity in the primate prefrontal cortex during associative learning. Neuron *21*, 1399–1407.

Asanuma, C., Andersen, R.A., and Cowan, W.M. (1985). The thalamic relations of the caudal inferior parietal lobule and the lateral prefrontal cortex in monkeys: divergent cortical projections from cell clusters in the medial pulvinar nucleus. J Comp Neurol *241*, 357–381.

Balan, P.F., and Gottlieb, J. (2006). Integration of exogenous input into a dynamic salience map revealed by perturbing attention. J Neurosci *26*, 9239–9249.

Balan, P.F., and Gottlieb, J. (2009). Functional significance of nonspatial information in monkey lateral intraparietal area. J Neurosci *29*, 8166–8176.

Balan, P.F., Oristaglio, J., Schneider, D.M., and Gottlieb, J. (2008). Neuronal correlates of the set-size effect in monkey lateral intraparietal area. PLoS Biol *6*, e158.

Barash, S., Bracewell, R.M., Fogassi, L., Gnadt, J.W., and Andersen, R.A. (1991a). Saccade-related activity in the lateral intraparietal area. I. Temporal properties; comparison with area 7a. J Neurophysiol *66*, 1095–1108.

Barash, S., Bracewell, R.M., Fogassi, L., Gnadt, J.W., and Andersen, R.A. (1991b). Saccade-related activity in the lateral intraparietal area. II. Spatial properties. J Neurophysiol *66*, 1109–1124.

Bichot, N.P., Rossi, A.F., and Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. Science *308*, 529–534.

Bisley, J.W., and Goldberg, M.E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. Science *299*, 81–86.

Bisley, J.W., and Goldberg, M.E. (2010). Attention, intention, and priority in the parietal lobe. Annu Rev Neurosci *33*, 1–21.

Brainard, D.H. (1997). The Psychophysics Toolbox. Spat Vis *10*, 433–436.

Braitenberg, V. (2007). Brain. Scholarpedia *2*, 2918.

Bruce, C.J., and Goldberg, M.E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. J Neurophysiol *53*, 603–635.

Bruce, V., and Young, A. (1986). Understanding face recognition. Br J Psychol *77 (Pt 3)*, 305–327.

Buffalo, E.A., Fries, P., Landman, R., Liang, H., and Desimone, R. (2010). A backward progression of attentional effects in the ventral stream. Proc Natl Acad Sci USA *107*, 361–365.

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*, 1860–1862.

Chafee, M.V., and Goldman-Rakic, P.S. (2000). Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades. J Neurophysiol *83*, 1550–1566.

Colby, C.L., and Goldberg, M.E. (1999). Space and attention in parietal cortex. Annu Rev Neurosci *22*, 319–349.

Colby, C.L., Duhamel, J.-R., and Goldberg, M.E. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. J Neurophysiol *76*, 2841–2852.

Conway, B.R., Moeller, S., and Tsao, D.Y. (2007). Specialized color modules in

macaque extrastriate cortex. Neuron *56*, 560–573.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. Machine Learning.

Dagnino, B., Gariel-Mathis, M.-A., and Roelfsema, P.R. (2015). Microstimulation of area V4 has little effect on spatial attention and on perception of phosphenes evoked in area V1. J Neurophysiol *113*, 730–739.

Darwin, C. (1872). The expression of the emotions in man and animals (London: J. Murray).

De Winter, F.-L., Zhu, Q., Van den Stock, J., Nelissen, K., Peeters, R., de Gelder, B., Vanduffel, W., and Vandenbulcke, M. (2015). Lateralization for dynamic facial expressions in human superior temporal sulcus. NeuroImage *106*, 340–352.

Deubel, H., and Schneider, W.X. (1996). Saccade target selection and object recognition: evidence for a common attentional mechanism. Vision Res *36*, 1827–1837.

Deubel, H., and Schneider, W.X. (2003). Delayed saccades, but not delayed manual aiming movements, require visual attention shifts. Ann N Y Acad Sci *1004*, 289–296.

Devlin, J.T., Russell, R.P., Davis, M.H., Price, C.J., Wilson, J., Moss, H.E., Matthews, P.M., and Tyler, L.K. (2000). Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. NeuroImage *11*, 589–600.

Durand, J.-B., Peeters, R., Norman, J.F., Todd, J.T., and Orban, G.A. (2009). Parietal regions processing visual 3D shape extracted from disparity. NeuroImage *46*, 1114–1126.

Ekstrom, L.B., Roelfsema, P.R., Arsenault, J.T., Bonmassar, G., and Vanduffel, W. (2008). Bottom-up dependent gating of frontal signals in early visual cortex. Science *321*, 414–417.

Fecteau, J.H., and Munoz, D.P. (2006). Salience, relevance, and firing: a priority map for target selection. Trends Cogn Sci (Regul Ed) *10*, 382–390.

Fisher, C., and Freiwald, W.A. (2015). Contrasting Specializations for Facial Motion within the Macaque Face-Processing System. Curr Biol *25*, 261–266.

Fox, C.J., Iaria, G., and Barton, J.J.S. (2009). Defining the face processing network: optimization of the functional localizer in fMRI. Hum Brain Mapp *30*, 1637–1651.

Freiwald, W.A., and Tsao, D.Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science

330, 845–851.

Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol *61*, 331–349.

Furl, N., Hadj-Bouziane, F., Liu, N., Averbeck, B.B., and Ungerleider, L.G. (2012). Dynamic and static facial expressions decoded from motion-sensitive areas in the macaque monkey. J Neurosci *32*, 15952–15962.

Ganguli, S., Bisley, J.W., Roitman, J.D., Shadlen, M.N., Goldberg, M.E., and Miller, K.D. (2008). One-dimensional dynamics of attention and decision making in LIP. Neuron *58*, 15–25.

Glimcher, P.W., and Sparks, D.L. (1992). Movement selection in advance of action in the superior colliculus. Nature *355*, 542–545.

Gnadt, J.W., and Andersen, R.A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. Exp Brain Res *70*, 216–220.

Gobbini, M.I., Gentili, C., Ricciardi, E., Bellucci, C., Salvini, P., Laschi, C., Guazzelli, M., and Pietrini, P. (2011). Distinct neural systems involved in agency and animacy detection. J Cogn Neurosci *23*, 1911–1920.

Gold, J.I., and Shadlen, M.N. (2007). The Neural Basis of Decision Making. Annu Rev Neurosci.

Goldberg, M.E., Gee, A., Ipata, A., Bisley, J.W., and Gottlieb, J. (2012). Attention and the Parietal Lobe. In The Neuroscience of Attention, G.R. Mangun, ed. (Oxford University Press), p. 167.

Goodale, M.A., and Milner, A.D. (1992). Separate visual pathways for perception and action. Trends Neurosci *15*, 20–25.

Gottlieb, J., and Goldberg, M.E. (1999). Activity of neurons in the lateral intraparietal area of the monkey during an antisaccade task. Nat Neurosci *2*, 906–912.

Gottlieb, J.P., Kusunoki, M., and Goldberg, M.E. (1998). The representation of visual salience in monkey parietal cortex. Nature *391*, 481–484.

Gottlieb, J., and Snyder, L.H. (2010). Spatial and non-spatial functions of the parietal cortex. Curr Opin Neurobiol *20*, 731–740.

Gregoriou, G.G., Gotts, S.J., and Desimone, R. (2012). Cell-Type-Specific Synchronization of Neural Activity in FEF with V4 during Attention. Neuron *73*, 581–594.

Gschwind, M., Pourtois, G., Schwartz, S., Van de Ville, D., and Vuilleumier, P. (2012). White-Matter Connectivity between Face-Responsive Regions in the Human Brain. Cereb Cortex.

Hanks, T.D., Kopec, C.D., Brunton, B.W., Duan, C.A., Erlich, J.C., and Brody, C.D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature.

Haxby, J., Hoffman, E., and Gobbini, M. (2000). The distributed human neural system for face perception. Trends Cogn Sci (Regul Ed) *4*, 223–233.

Helmholtz, von, H. (1867). Treatise on Physiological Optics Vol. III

.

Hikosaka, K. (1997). Responsiveness of neurons in the posterior inferotemporal cortex to visual patterns in the macaque monkey. Behav Brain Res *89*, 275–283.

Hikosaka, K. (1998). Representation of foveal visual fields in the ventral bank of the superior temporal sulcus in the posterior inferotemporal cortex of the macaque monkey. Behav Brain Res *96*, 101–113.

Hoffman, J.E., and Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. Percept Psychophys *57*, 787–795.

Huk, A.C., and Meister, M.L.R. (2012). Neural correlates and neural computations in posterior parietal cortex during perceptual decision-making. Frontiers in Integrative Neuroscience *6*, 86.

Ipata, A.E., Gee, A.L., and Goldberg, M.E. (2012). Feature attention evokes task-specific pattern selectivity in V4 neurons. Proceedings of the National Academy of Sciences *109*, 16778–16785.

Ipata, A.E., Gee, A.L., Bisley, J.W., and Goldberg, M.E. (2009). Neurons in the lateral intraparietal area create a priority map by the combination of disparate signals. Exp Brain Res *192*, 479–488.

Ipata, A.E., Gee, A.L., Goldberg, M.E., and Bisley, J.W. (2006). Activity in the lateral intraparietal area predicts the goal and latency of saccades in a free-viewing visual search task. J Neurosci *26*, 3656–3661.

Itti, L., and Koch, C. (2001). Computational modelling of visual attention. Nat Rev Neurosci *2*, 194–203.

Jastorff, J., and Orban, G.A. (2009). Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. J Neurosci *29*, 7315–7329.

Jastorff, J., Popivanov, I.D., Vogels, R., Vanduffel, W., and Orban, G.A. (2012). Integration of shape and motion cues in biological motion processing in the monkey STS. NeuroImage *60*, 911–921.

Juan, C.-H., Muggleton, N.G., Tzeng, O.J.L., Hung, D.L., Cowey, A., and Walsh, V. (2008). Segregation of visual selection and saccades in human frontal eye fields. Cereb Cortex *18*, 2410–2415.

Juan, C.-H., Shorter-Jacobi, S.M., and Schall, J.D. (2004). Dissociation of spatial attention and saccade preparation. Proc Natl Acad Sci USA *101*, 15541–15544.

Kanwisher, N., Mcdermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci *17*, 4302–4311.

Kiani, R., Cueva, C.J., Reppas, J.B., Peixoto, D., Ryu, S.I., and Newsome, W.T. (2015). Natural Grouping of Neural Responses Reveals Spatially Segregated Clusters in Prearcuate Cortex. Neuron.

Knudsen, E.I. (2012). Midbrain and forebrain systems for bottom-up control of spatial attention. In The Neuroscience of Attention, G.R. Mangun, ed. (New York: Oxford University Press), p. 131.

Kolster, H., Janssens, T., Orban, G.A., and Vanduffel, W. (2014). The Retinotopic Organization of Macaque Occipitotemporal Cortex Anterior to V4 and Caudoventral to the Middle Temporal (MT) Cluster. Journal of Neuroscience *34*, 10168–10191.

Konen, C.S., and Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. Nat Neurosci *11*, 224–231.

Kowler, E., Anderson, E., Dosher, B., and Blaser, E. (1995). The role of attention in the programming of saccades. Vision Res *35*, 1897–1916.

Kravitz, D.J., Saleem, K.S., Baker, C.I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. Nat Rev Neurosci *12*, 217–230.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., and Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends Cogn Sci (Regul Ed) *17*, 26–49.

Kriegeskorte, N., Formisano, E., Sorger, B., and Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. Proceedings of the National Academy of Sciences *104*, 20600–20605.

Ku, S.-P., Tolias, A.S., Logothetis, N.K., and Goense, J. (2011). fMRI of the face-

processing network in the ventral temporal lobe of awake and anesthetized macaques. Neuron *70*, 352–362.

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

Lee, J., and Maunsell, J.H.R. (2009). A normalization model of attentional modulation of single unit responses. PLoS ONE *4*, e4651.

Luck, S.J., Chelazzi, L., Hillyard, S.A., and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J Neurophysiol *77*, 24–42.

Mansouri, F.A., Matsumoto, K., and Tanaka, K. (2006). Prefrontal cell activities related to monkeys' success and failure in adapting to rule changes in a Wisconsin Card Sorting Test analog. J Neurosci *26*, 2745–2756.

Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84.

Mazer, J.A., and Gallant, J.L. (2003). Goal-related activity in V4 during free viewing visual search. Evidence for a ventral stream visual salience map. Neuron *40*, 1241–1250.

McCarthy, G., Puce, A., Gore, J.C., and Allison, T. (1997). Face-specific processing in the human fusiform gyrus. J Cogn Neurosci *9*, 605–610.

Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2000). Intermodal selective attention in monkeys. I: distribution and timing of effects across visual areas. Cerebral Cortex (New York, N.Y. : 1991) *10*, 343–358.

Meister, M.L.R., Hennig, J.A., and Huk, A.C. (2013). Signal Multiplexing and Single-Neuron Computations in Lateral Intraparietal Area During Decision-Making. Journal of Neuroscience *33*, 2254–2267.

Merigan, W.H. (2000). Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. Vis. Neurosci. *17*, 949–958.

Meyers, E.M. (2013). The neural decoding toolbox. Front Neuroinform *7*, 8.

Moeller, S., Freiwald, W.A., and Tsao, D.Y. (2008). Patches with links: a unified system for processing faces in the macaque temporal lobe. Science *320*, 1355–1359.

Moore, T., and Armstrong, K.M. (2003). Selective gating of visual signals by

microstimulation of frontal cortex. Nature *421*, 370–373.

Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. Science *229*, 782–784.

Mruczek, R.E.B., Loga, von, I.S., and Kastner, S. (2013). The representation of tool and non-tool object information in the human intraparietal sulcus. J Neurophysiol *109*, 2883–2896.

Neal, J.W., Pearson, R.C., and Powell, T.P. (1988). The organization of the cortico-cortical connections between the walls of the lower part of the superior temporal sulcus and the inferior parietal lobule in the monkey. Brain Res. *438*, 351–356.

Nelissen, K., Vanduffel, W., and Orban, G.A. (2006). Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. J Neurosci *26*, 5929–5947.

O'Toole, A.J., Roark, D.A., and Abdi, H. (2002). Recognizing moving faces: a psychological and neural synthesis. Trends Cogn Sci (Regul Ed) *6*, 261–266.

Ohayon, S., and Tsao, D.Y. (2012). MR-guided stereotactic navigation. Journal of Neuroscience Methods *204*, 389–397.

Oristaglio, J., Schneider, D.M., Balan, P.F., and Gottlieb, J. (2006). Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. J Neurosci *26*, 8310–8319.

Peck, C.J., Jangraw, D.C., Suzuki, M., Efem, R., and Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. J Neurosci *29*, 11182–11191.

Pinsk, M.A., Arcaro, M., Weiner, K.S., Kalkus, J.F., Inati, S.J., Gross, C.G., and Kastner, S. (2009). Neural representations of faces and body parts in macaque and human cortex: a comparative FMRI study. J Neurophysiol *101*, 2581–2600.

Pitcher, D., Dilks, D.D., Saxe, R.R., Triantafyllou, C., and Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage *56*, 2356–2363.

Posner, M.I. (1980). Orienting of attention. Q J Exp Psychol *32*, 3–25.

Premereur, E., Vanduffel, W., and Janssen, P. (2011). Functional heterogeneity of macaque lateral intraparietal neurons. J Neurosci *31*, 12307–12317.

Rajimehr, R., Young, J.C., and Tootell, R.B.H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. Proc Natl Acad Sci USA

106, 1995–2000.

Raposo, D., Kaufman, M.T., and Churchland, A.K. (2014). A category-free neural population supports evolving demands during decision-making. Nat Neurosci *17*, 1784–1792.

Rensink, R.A. (2002). Change detection. Annu Rev Psychol *53*, 245–277.

Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. Neuron *61*, 168–185.

Rigotti, M., Barak, O., Warden, M.R., Wang, X.-J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585–590.

Rishel, C.A., Huang, G., and Freedman, D.J. (2013). Independent category and spatial encoding in parietal cortex. Neuron *77*, 969–979.

Rizzolatti, G., Riggio, L., Dascola, I., and Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. Neuropsychologia *25*, 31–40.

Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci *22*, 9475–9489.

Rorie, A.E., Gao, J., McClelland, J.L., and Newsome, W.T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. PLoS ONE *5*, e9308.

Santella, A., and DeCarlo, D. (2004). Robust clustering of eye movement recordings for quantification of visual interest. Proceedings of the 2004 Symposium on Eye Tracking Research and Applications 27–34.

Sato, T.R., and Schall, J.D. (2003). Effects of stimulus-response compatibility on neural selection in frontal eye field. Neuron *38*, 637–648.

Schall, J.D., Morel, A., King, D.J., and Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams. Journal of Neuroscience *15*, 4464–4487.

Seltzer, B., and Pandya, D.N. (1980). Converging visual and somatic sensory cortical input to the intraparietal sulcus of the rhesus monkey. Brain Res. *192*, 339–351.

Sereno, A.B., and Maunsell, J.H. (1998). Shape selectivity in primate lateral intraparietal cortex. Nature *395*, 500–503.

Shadlen, M.N., and Gold, J.I. (2004). The neurophysiology of decision-making as a window on cognition. The Cognitive Neurosciences.

Smith, D.T., and Schenk, T. (2007). Enhanced probe discrimination at the location of a colour singleton. Exp Brain Res *181*, 367–375.

Smith, D.T., and Schenk, T. (2012). The Premotor theory of attention: time to move on? Neuropsychologia *50*, 1104–1114.

Snyder, L.H., Batista, A.P., and Andersen, R.A. (1997). Coding of intention in the posterior parietal cortex. Nature *386*, 167–170.

Sperling, G. (1960). The information available in brief visual presentations. Psychological Monographs: General and Applied *74*, 1.

Steinmetz, N.A., and Moore, T. (2014). Eye movement preparation modulates neuronal responses in area v4 when dissociated from attentional demands. Neuron *83*, 496–506.

Stemmann, H., and Freiwald, W.A. (2015). Cortical area PITd, a ventral pathway area for the control of spatial visual attention?

Stemmann, H., and Freiwald, W.A. (2015). An Area in Inferotemporal Cortex Supports Attentive Motion Processing in the Macaque. 1–39.

Suzuki, M., and Gottlieb, J. (2013). Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. Nat Neurosci *16*, 98–104.

Tsao, D.Y., Freiwald, W.A., Knutsen, T.A., Mandeville, J.B., and Tootell, R.B.H. (2003). Faces and objects in macaque cerebral cortex. Nat Neurosci *6*, 989–995.

Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H., and Livingstone, M.S. (2006). A cortical region consisting entirely of face-selective cells. Science *311*, 670–674.

Tsao, D.Y., Moeller, S., and Freiwald, W.A. (2008). Comparing face patch systems in macaques and humans. Proceedings of the National Academy of Sciences *105*, 19514–19519.

Tsotsos, J.K. (2011). A Computational Perspective on Visual Attention (MIT Press).

Ungerleider, L.G., and Mishkin, M. (1982). Two cortical visual systems. In Analysis of Visual Behavior, M.A. Goodale, D.J. Ingle, and R.J. Mansfield, eds. (Cambridge: MIT Press).

Ungerleider, L.G., and Pessoa, L. (2008). What and where pathways. Scholarpedia *3*, 5342.

Van Essen, D.C., Glasser, M.F., Dierker, D.L., and Harwell, J. (2011). Cortical Parcellations of the Macaque Monkey Analyzed on Surface-Based Atlases. Cereb Cortex.

Vanduffel, W., Fize, D., Mandeville, J.B., Nelissen, K., Van Hecke, P., Rosen, B.R., Tootell, R.B., and Orban, G.A. (2001). Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. Neuron *32*, 565–577.

Ward, L.M. (2008). Attention. Scholarpedia *3*, 1538.

Yovel, G., and Kanwisher, N. (2004). Face perception: domain specific, not process specific. Neuron *44*, 889–898.

Zénon, A., and Krauzlis, R.J. (2012). Attention deficits without cortical neuronal deficits. Nature *489*, 434–437.

Zhang, M., and Barash, S. (2000). Neuronal switching of sensorimotor transformations for antisaccades. Nature *408*, 971–975.