Neural Basis of Executive Control in Prefrontal Cortical Networks

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Shikha Jain Goodwin

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Matthew V. Chafee
Co-Advisor: James Ashe

April, 2012

Acknowledgements

I want to thank my parents who supported me. My dad bought books, provided me with the resources for my education and taught me the ideology of giving studies the top most priority in life.

The most important person in helping me finish my dissertation has been my husband Tom. He has been my pillar and my strength. It is suffice to say I wouldn't have been able to complete this process without his support.

Art has been one person in my life who has helped me overcome a variety of struggles, both professionally and personally during my PhD process. I am extremely grateful to him.

I want to extend my thanks to a number of good friends- Rakhi, Sarah, and many more. I would run out of room, if I start naming them all. My friends helped make my life easier during this phase of my life.

I would like to thank my thesis committee – Dr. Georgoupolos, Dr. Jim Ashe and Dr. Tay Netoff- who helped and guided me with my research. There are a number of people in the labs of Dr. Chafee and Dr. Georgoupolos who have provided me with professional support – Dale, Dean, Dave, Vijay and many others.

My greatest appreciation is for Matt- who had been like a father to me. His support, kind words, invaluable advice and more importantly being there for me has been invaluable to get through this process.

Last, but not the least I would like to thank God for providing this wonderful opportunity and giving me patience and determination.

Dedication

This dissertation is dedicated to my dad who made it possible for me to reach my lifelong dream of having a PhD and to my husband Tom who made sure that I survived the strenuous process.

Abstract

The goal of this research is to better understand how prefrontal cortical networks mediate executive control, the capacity to select and implement different information processing algorithms as a function of goals, rules, or other internal state variables. We trained two rhesus macaque monkeys to place visual stimuli into spatial categories according to a variable rule. We refer to the behavioral paradigm we developed for this purpose as the Dynamic Spatial Categorization (or DYSC) task, because the task requires the brain to classify each stimulus differently according the rule or grouping principle that is in force. Behavioral performance provided evidence that the monkey placed visual stimuli into spatial categories according to the variable rule. We then recorded neural activity from prefrontal and posterior parietal cortex simultaneously during task performance using dual, depth-adjustable 16 microelectrode arrays from ensembles of neurons. We found that the neural representation of the category and executive control of category as evidenced by the rule dependence of that representation were distributed in the network. The neural signal coding spatial category exhibited rule-dependence in prefrontal cortex first however, consistent with prefrontal cortex leading in the network implementation of executive control.

Contents

Chapter 2 Executive control of spatial cognition in prefrontal-parietal network

List of Figures

INTRODUCTION

Executive control of cognitive processing

This research will examine the neural basis of executive control in the brain. Executive control refers to the brain's ability to switch between competing cognitive processes and information processing operations as a function of internal goals. This selection can be internally driven and controlled and is not necessarily triggered by or dependent on a change in the pattern of sensory input. The ability of the brain to select and control which cognitive process to apply to the sensory input in order to select an action as a function of the goal at hand, enables it to change which actions are taken in response to any given stimulus. When one stimulus is not necessarily associated with one particular response, then it is often necessary to respond to that stimulus differently as a function of different goals or contexts. This is the essence of executive control. Thus, executive control is the ability of the brain to control, plan, strategize, pay attention to and execute different cognitive and neural processes to achieve specific goals. Executive control allows the brain to extract the most important information from the sensory input and process it in a context dependent fashion to produce the desired action(Stoet and Snyder 2006), demonstrating the computational flexibility that typifies human cognition and behavior. My objective in this dissertation is to elucidate executive control processes at a cellular level in cortical networks and to understand how systems of interacting neurons implement executive control in the brain. This will give us mechanistic insight into an important dimension of human intelligence (Goldman-Rakic

1995; Goldman-Rakic 1996; Miller and Cohen 2001; Freedman and Miller 2008; Tanji and Hoshi 2008).

Role of prefrontal cortex in executive control

There is considerable evidence that executive control depends on the functional integrity of prefrontal cortex. One line of evidence (discussed at greater length below) is based on the loss of cognitive flexibility and the ability to use rules to direct behavior in both monkeys and humans following damage to prefrontal cortex (Goldman-Rakic 1995; Goldman-Rakic 1996; Freedman and Miller 2008). These deficits can be attributed to a loss either in the ability of the brain to generate and maintain internal representations of behavioral goals or rules or to switch between them. Executive control requires working memory, as the rule needs to be buffered and maintained in memory before it can be applied to incoming sensory information to generate contextually appropriate responses. Most cognitive processes and behaviors that are under executive control involve multiple steps that are extended in time, making working memory essential. Thus, it has been hypothesized that working memory is a key component of executive control (Goldman-Rakic 1995; Goldman-Rakic 1996), inhibition of prepotent responses, and the learning and representation of rules and category information (Fuster, 2001; Miller and Cohen, 2001).

Prefrontal lesions and dysexecutive syndrome The Case of Phineas Gage (Harlow 1868) is a very famous case in history in which a man suffered damage to prefrontal cortex by a

tamping iron blown through his skull (specifically through the left prefrontal cortex). After the injury, Gage had an excellent physical recovery but there was a dramatic personality change. People described him as 'no longer Gage' with the appearance of new personality characteristics such as being stubborn, lacking in consideration for others, having profane speech, and failing to execute his plans. His physician at the time, Dr. Harlow, described that "his equilibrium, or balance, so to speak, between his intellectual faculties and animal propensities seems to have been destroyed." Also, Gage was no longer able to go to his job or able to organize his behavior, indicating deficits in executive processes.

The effects of PFC lesions are similar in several animal species, and to elucidate the effects of lesions, we will discuss the results of various studies performed in the last century. Hitzig (1874) and Ferrier (1886) described 100 years ago the effects of experimental ablations of PFC in dogs concluding that the region is particularly involved with intellectual functions and attention rather than sensory and motor functions. They summarized their findings by noting "the loss of ability to excite the motions necessary for carrying out the will." In 1922, the Italian physiologist Bianchi concluded that removal of frontal lobes in non-human primates destroyed the ability of animal to synthesize various incoming percepts and integrate them with outgoing motor commands. Jacobsen's pivotal work in 1936 demonstrated that the monkeys whose prefrontal cortex is removed showed a selective and delay dependent deficit on delayed response task. Monkeys would observe an experimenter place a food reward in one of two wells, and

then cover the wells with plaque, and their task was to chose the location of food well. After experimental lesions of the prefrontal cortex, monkeys were at near chance performance in retrieving the bait, demonstrating a deficit in spatial working memory.

Jacobsen CF (1938). "Studies of cerebral function in primates". *Comp Psychol Monogr* 13: 1–68.

The involvement of prefrontal cortex in executive control is further supported by the fact that patients with damage to prefrontal cortex have difficulty using rules to direct behavior. Petrides found, for example, that following PFC damage, patients could no longer learn associations between visual patterns and hand gestures (Petrides 1985). These tasks required the learning of arbitrary associations between a set of stimuli and a set of responses. Patients with excisions from the left or right frontal cortex were severely impaired in learning both tasks. Executive control deficits are also apparent when tested with standardized neuropsychological tests. The Wisconsin Card Sorting Test (WCST) for example is a neuropsychological test of "set-shifting", i.e. the ability to display flexibility in cognitive processing as required by a changing rule within a task. It is a test that is sensitive to prefrontal damage. In this task, subjects sort cards containing geometric shapes in different colors and numbers according to a changing sorting rule (based on color, shape or number). Subjects are required to discover the rule by trial and error. For example, in the first trial, if they pick the rule to be “shape” and it’s wrong, it gives them the indication to switch the rule and apply either “color” or “number” as a

grouping criterion in a subsequent trial. When subjects get the rule right, they must continue to use the rule on subsequent trials, until the rule changes again (without notification by the experimenter) and they have to repeat the whole process to find the new correct rule. Patients with damage to frontal lobe do poorly on this test. Particularly, in comparison to control subjects, patients with dorsolateral frontal lobe lesions make a higher number of preservative errors (errors in which subjects continue to sort according to a rule they have been informed is incorrect) instead of adapting and switching to the new rule (Goldman-Rakic, Bates et al. 1992; Greve, Stickle et al. 2005).

Stuss summarized the results of neuropsychological testing in patients with frontal lobe lesions to conclude that in these cases, there is a lack of 'supervisory' processes that are important for control of lower-order processes (Stuss, 2007). He examined the results from four different classic 'frontal' tests used to determine deficits in executive processing including WCST; Stroop test (Stroop1935; Comalli et al. 1962); language-verbal fluency (Borkowski et al. 1967; memory—list learning (Delis et al. 1987); ROBBIA (ROtman-Baycrest Battery to Investigate Attention). In the stroop task, there are words describing various colors written with different inks. Subjects are asked to report the color of ink of a word. There are two cases, congruent in the case that the written word spells out the same as color as the color of the letters themselves and an incongruent condition, in which the written word indicates a different color than that of the color of the letters. The task is for subjects to report the color of the letters, and ignore the color indicated by the word. One effect noted with prefrontal lesion patients is that the

reaction time is longer on incongruent trials in comparison to control group, interpreted as evidence that prefrontal patients have to exert greater effort to override the habitual response of reading the word. In addition, after prefrontal damage, patients more than controls report the color indicated by the word, and fail to inhibit reading altogether. ROBBIA tests consisted of 2 variations : 1) Simple RT - the capital letter 'A' was presented 50 times repeatedly and subjects responded by pressing button 1 as soon as the letter was presented. And 2) The target was defined as one of the four letters (A, B, C and D) presented 25% of the time, the remaining letters being non-target feature distracters, each also presented with a probability of 25%. Responses were made to both targets and non-targets. After a block of trials, the target and distracter letters were switched requiring a remapping of the motor response. It has been demonstrated that the patients with lesions have critical impairment in performing the task and separating targets and non-targets, providing evidence that they have difficulty maintaining attention for response control.

Neural correlates of executive control in prefrontal cortex: human studies Functional imaging studies have provided further evidence of PFC recruitment to mediate executive control. Miyashita and colleagues trained monkeys to perform a simplified version of the WCST, and compared the activity of monkeys and humans performing the same task. By contrasting BOLD **(**blood-oxygen-level-dependent**)** fMRI activity between trials that involved switching versus the trials in which the rule was repeated (Nakahara, Hayashi et al. 2002). They identified a region in inferior prefrontal cortex that was more strongly

activated on switch trials (trials that involve changing the rule) in both monkeys and humans. This provided evidence that prefrontal cortex was particularly engaged when the brain has to apply a new rule to control behavior, rather than repeat the previously rewarded rule or response. The study also found that prefrontal activation was greater when working memory was required to maintain the rule (Hayashi, Konishi et al. 1999), providing evidence of the link between working memory and executive control. These studies indicate that the prefrontal cortex plays an essential role in executive control. However, how prefrontal cortex implements its role in executive control by interacting with neurons in other cortical areas remains relatively poorly understood.

Neural correlates of executive control in prefrontal cortex: nonhuman primate studies. To review current understanding of the neural mechanisms of executive control in PFC at the cellular level, I will focus on single neuron studies in non-human primates. An example of how prefrontal neurons are engaged by executive control processing in rule-based processing was provided by Assad et al. (Asaad, Rainer et al. 2000), who trained monkeys to alternate between tasks that employed the same cues and responses, but different rules governed the relationship between cues and responses. Monkeys viewed visual stimuli and selected saccade targets according to three different rules: matching (delayed matching to sample), associative (conditional visuomotor), and spatial (spatial delayed response). Over half of lateral PFC neurons were rule dependent. Neural responses to a given cue or forthcoming saccade often depended on which rule was current in the task. In addition, the baseline activity of many neurons (54%) varied with

the rule itself. Hoshi et al. have also observed PFC neurons with activity that was modulated by whether the monkey was using a shape-matching or location-matching rule (Hoshi, Shima et al. 1998). Further, Roberts et al. have shown that lateral and orbito-frontal PFC neurons reflect whether the monkey is currently using a "matching" or "nonmatching" rule to select a test object (Roberts and Wallis 2000).

In their recent paper Wise et al. found that activity correlated with the strategy used for choosing the goal in the population of prefrontal cortex neurons that encoded either a previous goal or a future goal, chosen by monkeys according to specific behavioral strategies. The task included monkeys making saccadic eye movements to different targets according to different abstract response strategies signified by a visual cue (Tsujimoto, Genovesio et al. 2008). The strategies were either repeat-stay (repeat the last motor response if the same cue appeared twice in a row), or change-shift (if a new cue was presented, select a new target). These authors found neurons in prefrontal cortex that were selectively activated for each of the two alternative strategies. Tanaka et. al. found that the dorsolateral prefrontal cortex is activated in maintaining the relevant rule across trials during the WCST task for example, whether the monkey was matching according to number, color, or shape (Mansouri, Matsumoto et al. 2006). These data indicate that the prefrontal cortex is important in updating, maintaining, or switching set; as required by the changing rule of the WCST task.

There have been studies in the past that have referenced task related and rule related activity. Assad and colleagues trained monkeys to alternate between tasks that employed the same cues and responses but three different rules: matching (delayed matching to sample), associative (conditional visuomotor), and spatial (spatial delayed response), and found that over half of lateral PFC neurons were rule dependent. Neural responses to a given cue or forthcoming saccade often depended on which rule was current in the task. In addition, the baseline activity of many neurons (54%) varied with the rule (Asaad, Rainer et al. 2000). In this study though, neural recordings were restricted to prefrontal cortex, rules were not explicitly instructed and the task was sensorimotor in nature and did not require abstract neural representations (such as computation of category membership). Thus, we need to study the neural mechanisms by which rules control the generation of abstract representations at the network level; to demonstrate how the prefrontal cortex influences other cortical areas to mediate executive over cognitive (as opposed to sensorimotor) processing.

Relation of executive control to working memory Working memory refers to the process of temporarily storing and manipulating information in the brain. Research to date (Goldman-Rakic 1995; Goldman-Rakic 1996) has established the fact that the prefrontal cortex plays an essential role in supporting working memory. Working memory is essential to executive control in many circumstances because whatever goal or rule is selected to direct behavior, it has to be maintained in working memory for the period of time that behavior is controlled by that rule. In the nonhuman primate, working memory

has typically been studied using the delayed response tasks. The Spatial Delayed Response Task (SDR) has been used extensively for assessing spatial working memory. Spatial working memory refers to the temporary retention of spatial information in a short-term buffer that can then be manipulated to support spatial cognition or utilized to direct future action. In a delayed response task, monkeys view a brief visual cue and then must remember that stimulus location for a brief interval after it disappears in order to successfully direct a future response. The oculomotor delayed response task (ODR) is an example of a working memory task in which the motor response is a saccadic eye movement. In the ODR task, subjects (either humans or monkeys) focus their gaze on a small fixation target presented at the center of a video display. While fixating, a second small circular cue flashes in one of several (typically 8) peripheral locations, followed by a delay period of several seconds (typically 3). After the delay, subjects are required to execute a memory-guided saccade to the spatial location where they remember the cue that had just appeared. Neural recording experiments in monkeys performing the ODR task revealed that the prefrontal neurons exhibited a pattern of electrical activity that seemed to implement working memory at a cellular level. These neurons were activated by the presentation of the visual cue, were selective for its location in space, and sustained this elevated activity after the cue disappeared throughout the subsequent delay period. Funahashi et al. found that many neurons in the dorsolateral prefrontal cortex, including and surrounding the principal sulcus and in the frontal eye field, exhibited mnemonic persistent activity during the delay period (Funahashi, Bruce et al. 1989). On the basis of the observation that all spatial locations tested in the experiment were

represented across the recorded neurons, the authors concluded that different prefrontal cells encode and store in working memory different spatial locations so that "this area of the cortex contains a complete 'memory' map of visual space". The stored spatial information provides control of the behavior in the absence of instructive sensory input at the time of the response. This property of working memory has been recognized as an essential aspect of executive control (Goldman-Rakic 1995; Goldman-Rakic 1996).

Anatomical and functional properties of prefrontal cortical networks

Prefrontal connectivity Prefrontal cortex is densely connected with other areas and its contribution to cognition is likely to depend on this interaction. Our research focuses on the parietal-prefrontal network, which is one node of a distributed network of connections that link prefrontal cortex to other higher order cortical association areas. Figure 1 shows the extrinsic and intrinsic connections of the prefrontal cortex. These pathways were identified by injecting axonal tracers into the prefrontal cortex and other areas and visualizing the input and output projections that linked them into complex networks. Inputs to the prefrontal cortex consist of connections from the extrastriate visual areas, posterior parietal cortex, inferotemporal cortex, auditory cortex, amygdala, hypothalamus, hippocampus, thalamus and limbic areas.

Most of the connections between prefrontal cortex and other cortical areas are reciprocal (Cavada and Goldman-Rakic 1989). PFC can influence action by output

projections to a number of motor areas, including premotor cortex, the supplementary motor area, the caudate nucleus, and the cerebellum (Cavada and Goldman-Rakic 1989; Cavada and Goldman-Rakic 1989). There are cingulate motor areas on the medial wall of the PFC that have neurons that project to the spinal cord and are active in relation to motor control.

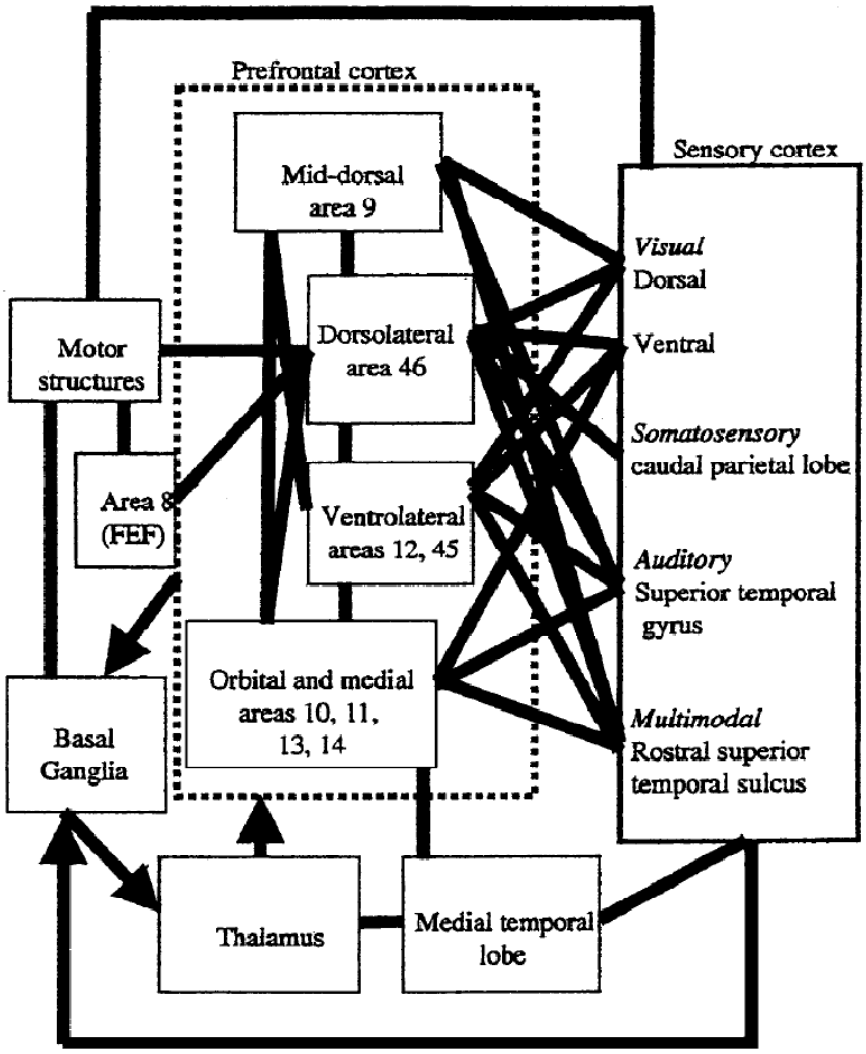

Figure 1. Connections to and from Prefrontal Cortex (Miller and Cohen 2001)

PFC is therefore perfectly suited to monitor neural activity across a wide array of sensory and motor processing systems and exert a supervisory or executive influence over our behavior, consistent with the view that PFC is involved in guiding and controlling complex goal directed behavior in both humans and nonhuman primates (Freedman and Miller 2008). This diversity of connections and exchange of signals between PFC neurons and other areas is likely the reason why PFC neurons are (1) activated by diverse stimuli, (2) are active prior to movements, (3) can carry information about internal factors such as motivation and attention.

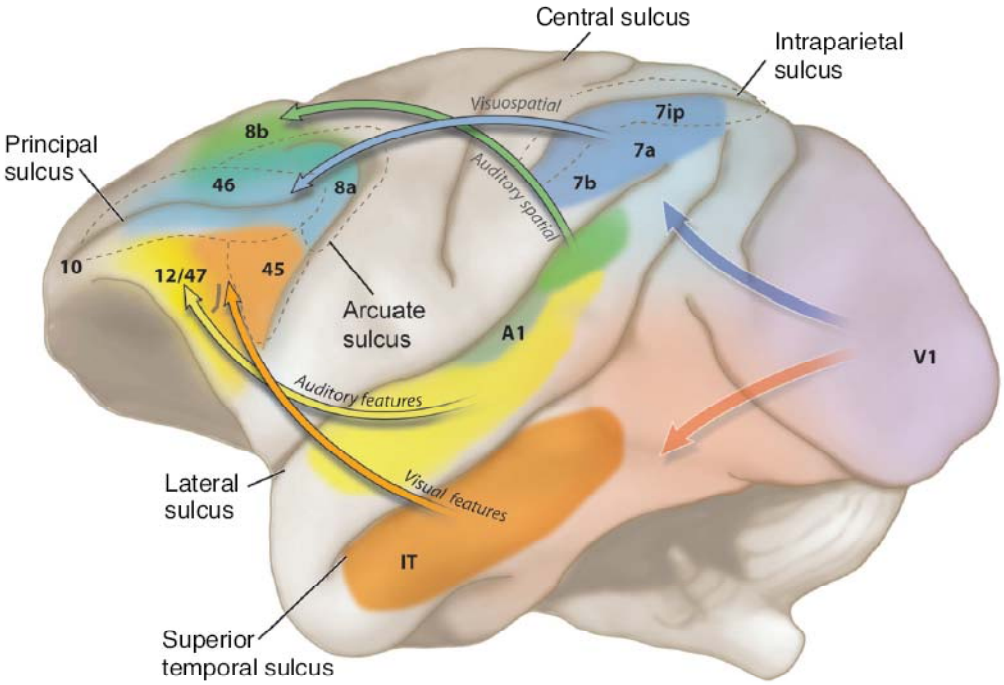

Figure 2. Summary of PFC input projections from sensory association cortices in parietal and temporal cortex to the prefrontal cortex of nonhuman primates (Cavada and Goldman-Rakic 1989)

A Distributed System: The Parietal-Prefrontal Network

The cerebral cortex is functionally and anatomically organized into distributed systems, which are defined as groups of cortical areas that are connected to each other through axonal projections which help to transfer neural signals from one part of the brain to another (Mountcastle 1997). PFC sits at the center of a number of distributed systems linking sensory systems of different modality to motor areas.

One such network appears to be involved primarily in spatial cognitive function, and involves the reciprocal pathway between parietal and prefrontal cortex. These areas are also connected to each other anatomically and are functionally related (Cavada and Goldman-Rakic 1989). Prefrontal neurons give rise to axons that synapse directly onto parietal neurons and parietal neurons give rise to axons that synapse directly on prefrontal neurons. Each subdivision of parietal cortex is connected with a unique set of frontal areas. My thesis research focuses on one component of the parietal-prefrontal network that consists of neurons in posterior parietal areas 7a in the inferior parietal gyrus, and neurons in prefrontal areas 46 in and around the principal sulcus. Neurons in these subdivisions of prefrontal and posterior parietal cortex project directly to one another (Cavada and Goldman-Rakic 1989)

Functionally the two areas, prefrontal and parietal, exhibit similar patterns of activity during behavior. For example, Chafee and Goldman Rakic (1998) found matching patterns of activity in another pair of reciprocally connected prefrontal and

parietal areas, prefrontal area 8a and parietal area LIP during a spatial working memory task (Chafee and Goldman-Rakic 1998). Monkeys had to remember the location of a peripheral cue in order to direct a memory guided saccade to that location 3 seconds after the cue disappeared. Chafee and Goldman-Rakic found that the visual, working memory, and motor signals in prefrontal and parietal cortex were nearly identical. This suggests that distributed systems share signals during information processing and that neural representations of task related information takes place at a network level rather than the level of individual cortical areas. To test the hypothesis that the reciprocal projections between parietal and prefrontal neurons may entrain their parallel activation, Chafee and Goladman Rakic performed experiments in which they combined cortical cooling in one cortical area with single-unit recording in the other to more precisely determine the physiological interactions between the two during working memory performance (Chafee and Goldman-Rakic 2000). The authors found that ODR performance changed significantly in both cases. This suggests that prefrontal and parietal neurons achieve matched activation during ODR performance through a symmetrical exchange of neuronal signals between them and the output of each cortical area produces a mixture of excitatory and inhibitory drives within its target.

To test the possibility that prefrontal cortex might exert executive control over parietal cortex during spatial cognition, we have designed and executed a series of experiments to investigate the executive control of spatial cognitive processing and will study how computational flexibility in spatial cognition is reflected in patterns of neural

activity within the prefrontal-parietal network. This study will address how rules control cognition, as opposed to sensorimotor processes. The experiments are designed to isolate neural signals associated with spatial cognitive computation (placing a stimulus into spatial category) and then to show how rules control this computation by modulation of the neural signals involved.

Spatial representation in posterior parietal cortex

Role of parietal cortex in visuomotor control The parietal-prefrontal network is a likely substrate for the executive control of spatial cognition because, just as considerable evidence implicates PFC in executive control (see above), comparable evidence implicates parietal cortex in spatial sensorimotor processing and spatial cognition. Posterior parietal cortex is known as an association area, as it combines information from different sensory modalities to form a cognitive representation of space. The posterior parietal cortex is functionally situated between sensory cortex and motor cortex. Thus, posterior parietal cortex is anatomically situated to act as an interface between sensory and motor structures and performs various functions important for sensory-motor integration. Buneo 2006 presents a review that considers the functional contribution of parietal cortex, characterizing it as a sensorimotor interface for visually guided movements (Buneo and Andersen 2006). Lesions in the parietal association cortex produce profound disturbances in visually guided behavior (Mountcastle, Lynch et al. 1975). The major conclusion from this study was that there exist, within the posterior parietal association cortex, sets of neurons which function as a command apparatus for

the behavioral acts of manual and visual exploration of extra personal space, directing movements of eyes and arms to explore the immediately surrounding area. Subsequent work by Anderson and colleagues supported a sensorimotor role for parietal cortex. By recording neural activity within parietal cortex as monkeys shifted gaze and viewed visual stimuli these authors found that single neurons in inferior parietal lobule in behaving monkeys possess retinotopic visual receptive fields, and that the sensitivity of these receptive fields changed systematically with the angle of gaze (Andersen, Essick et al. 1985). This joint sensitivity means that parietal neurons were maximally activated when a stimulus fell in a particular location on the retina, and when the eyes were directed in a particular direction, such that neurons were maximally activated for stimuli that were located at particular locations defined in relation to the head, a type of spatial representation referred to as craniocentric coding. Later studies demonstrated that the visual responses of parietal neurons were further modulated by rotation of the head relative to the body (Snyder, Batista et al. 2000) providing evidence that visual stimuli were coded in coordinates defined relative to the body. Head-centered or body-centered representations of space are generated by the combination of eye position, postural, and retinal factors. These representations may be useful for the visual guidance of movement because they define the locations of targets relative to the body. The involvement of parietal neurons in generating electrical signals that could serve as motor commands has been substantiated by other neural recording studies. Snyder and colleagues recorded in area LIP, found that activity varied as a function of the direction of a forthcoming saccade (Snyder, Batista et al. 2000). This activity could reflect the direction of attention

toward the saccade target, but these authors provided evidence that there was a strong modulation of LIP activity when movement plans were changed without changes in the locus of spatial attention (Snyder, Batista et al. 2000), supporting a purely motor role. Thus, the authors suggested that PPC, which has been postulated to play a role in shifting attention, may also play a role in changing movement intentions, or motor plans. These data have led to the view that parietal cortex plays a role in motor planning, helping to convert spatial representations from eye-centered to body-centered coordinates, thus implementing essential spatial transformations for visuomotor control.

Role of parietal cortex in spatial attention. As discussed above, Mountcastle and colleagues described cells in posterior parietal cortex of the rhesus monkey (area 7) that discharge in association with eye movements and hand movements (Mountcastle, Lynch et al. 1975). Goldberg et al. also studied the activity of cells in area 7 during visually guided saccadic eye movements, visual fixation, smooth-pursuit eye movements, and visually guided reaching movements (Robinson, Goldberg et al. 1978). Goldberg and colleagues suggested that the motor signals described by Mountcastle's group actually reflect the position of the visual stimuli used as the targets for movement, or the direction of attention to these stimuli. This started a very famous debate in 1970's. Mountcastle group argued that parietal cortex was activated in relation to motor intention, whilst Goldberg argued that parietal cortex activation was solely visual and/or attention based. This debate continues to date, and there are recent findings to support both claims. For example, Goldberg et al. found that, in addition to representing object locations in motor

coordinates, parietal neurons exhibit strong modulation by attention (Colby and Goldberg 1999) and that the spatial location coded by apparently motor planning related activity correlated with the position of visual attention (Bisley and Goldberg 2003). Thus, the relative roles of parietal cortex in intention and attention remain unresolved. It appears that spatial representation in parietal cortex influences both aspects of behavior, both attention and motor control.

Role of parietal cortex in spatial cognition Although visuomotor functions as reviewed above are well documented, parietal cortex is also involved in higher cognitive functions. For example, it has been shown that parietal neurons encode signals related to behavioral rules under which monkeys process visual stimuli and respond to them (Stoet and Snyder 2004). Crowe and colleagues have recently explored patterns of neural activity in posterior parietal cortex the context of various spatial cognitive tasks**.** In one of these, monkeys were required to determine from a single point of fixation whether a critical path through the maze reached an exit or a blind ending (Crowe, Chafee et al. 2004). Crowe et al. found that during this process, the activity of approximately one in four neurons in area 7a was spatially tuned to the direction of the maze path that monkeys mentally followed, without the monkey ever making a physical movement (of eyes or arm) based on this direction. This experiment concluded that the neural activity reflected a spatial cognitive analysis of the maze structure.

In another study, Chafee et al. recorded the activity of parietal area 7a neurons in monkeys performing an object construction task. In each trial, a model object consisting of a variable arrangement of squares was presented, followed after a delay by a copy of the model object that was missing a single square. Monkeys replaced the missing square to reconstruct the model configuration. Activity of many 7a neurons varied systematically with the position of the missing square and predicted where monkeys were going to add parts to the object they were building (Chafee, Crowe et al. 2005). In this respect, neural activity in parietal area 7a reflected spatial computations required to solve the task, and these signals were dissociated from sensorimotor factors (such as stimulus position or movement direction). Further, by shifting the position of the reference object that monkeys were constructing relative to the gaze fixation target, these authors were able to dissociate spatial signals defined in relation to the viewer from those defined in relation to the object. Most of the neurons activated to reflect the position of the missing square coded this location in object-centered coordinates, with activity varying as function of whether the missing square was located to the left or right of the reference object midline, regardless of where the missing square was located with respect to the viewer (Chafee, Averbeck et al. 2007).

Chafee et al. used time resolved ensemble decoding based on linear discriminate analysis (LDA) to measure short-term variation in the information about viewer-centered and object-centered spatial position coded by the activity of simultaneously recorded neural populations in parietal cortex (Chafee, Crowe et al. 2005). Object referenced

representations must be derived from retina or viewer referenced representations, as visual input is initially provided in a retina-based framework. Consistent with this, these investigators found that viewer-centered signals appear first in parietal cortex, followed by object-centered signals. They also found that object-centered representation was lagged and correlated in strength with viewer-referenced representation over time. This is consistent with the possibility that viewer-referenced signals provide input to the transform that generates object-referenced signals, and that this transform takes place in part within parietal cortex.

Studies of the role of parietal cortex in spatial cognition also demonstrated that spatial representations in parietal cortex can be highly context dependent. For example, Crowe et al. quantified the directional preference of single parietal neurons during maze solution, and found that preferred directions during maze solution bore little systematic relationship to the preferred directions exhibited by the same neurons in simpler processing contexts, such as when making visually-guided movements to visual targets (Crowe, Chafee et al. 2004). Their experimental results suggest that training on a specific task, such as the mental maze task, results in task specific representations and that these representations, do not necessarily generalize to other task contexts. Similarly, neurons that code positions in objects during object construction are generally inactive when the same locations serve as targets for saccadic eye movements or contain attended visual targets (Chafee, Crowe et al. 2005).

Tasks such as the object construction and maze tasks required monkeys to perform increasingly complex and rule dependent spatial analyses of the visual input, and parietal neurons generated spatial representations that were custom tailored to the computational demands of the task being performed (Chafee, Crowe et al. 2005; Crowe, Averbeck et al. 2005; Chafee, Averbeck et al. 2007), providing evidence that parietal cortex participated in spatial cognitive processing in addition to spatial processing for visuomotor control. Consistent with this role in higher spatial cognitive functions, recently it has been reported that neural activity in LIP reflects how a monkey categorizes a moving dot stimulus based on the direction of motion, and this categorical activity adapts to reflect new category boundaries with training (Freedman and Assad 2006), in a manner similar to the adaptation of category representation over learning that has been demonstrated in prefrontal cortex (Freedman, Riesenhuber et al. 2001). A recent study compared neural signals that reflected categorization of a visual stimulus as 'high' or 'low', and found that neural signals reflecting these categories were distributed between parietal and prefrontal cortex, although the signal emerged earlier in prefrontal cortex (Merchant, Crowe et al. 2011). Studies investigating the neural basis of spatial categorization in parietal cortex are particularly relevant to the current project, and are discussed in greater detail in the Introduction to Chapter 1.

A test case for the executive control of spatial cognition: Involvement of parietal cortex in spatial categorization

Likelihood that categorization depends on executive control. The above studies describing the anatomy and physiology of the prefrontal-parietal network suggests that interaction between these cortical areas is likely to be important for the executive control of spatial cognition. Categorization provides a good test case for the executive control of cognition, because the category to which a stimulus belongs is often a function of a rule. Any single object can belong to a large number of categories. And thus, the category it will belong to depends on the grouping criterion employed, and this in essence is a type of rule. Thus, any stimulus can be categorized according to innumerable criteria, and selection of one particular grouping or categorizing criterion amounts to the selection of a rule to control the cognitive process that categorizes, and this requires executive control. Since, our interest is in neural mechanisms of spatial categorization we decided to investigate neural activity in the prefrontal-parietal network to identify neural signals related to the executive control of this cognitive process. For the purposes of our experiments, we will define a spatial category as a region of space in which all points share a common spatial relationship to a category boundary. For example, given a line of particular orientation, one spatial category would contain all points to the left or above that line. In this way, spatial categorization is the product of the transformation of neural signals that code spatial position of a stimulus into signals that represent a category to which the stimulus is assigned, based the position of the stimulus relative to the boundary. Thus, categorization functions reflects the generation of a cognitive representation of sensory information that is the product of a transform applied to that

input, and we can record the neural signals by which the transformation is controlled and modified on a trial-by-trial basis.

Often, categorization is thought of as a process in which the brain imposes a boundary on a stimulus continuum, dividing it into discrete regions. Spatial categorization could provide unique insight into this process because the signals representing the input and output of the transform are relatively easy to quantify. It is easy to measure and control the spatial positions of stimuli, as well as the spatial position of boundaries that parse space into categories, and to dissociate neural signals that code these spatial variables from the spatial category of a stimulus. By independently manipulating the spatial position and spatial category of a stimulus, we can dissociate distinct neural signals coding these two variables. Most of the work on categorization in prefrontal cortex has focused on categorizing the form of visual objects (Freedman, Riesenhuber et al. 2001). In Freedman 2001, they trained monkeys to characterize visual objects as cats or dogs. The images of dogs and cats were morphed together in various combinations and the monkeys used groups of distinctive features to classify each image as being a dog or a cat. The gradual morphing between dog and cat prototypes produced a stimulus continuum on which the monkeys imposed a category boundary, based on the pattern of their behavioral report. They found that the activity of prefrontal cortical neurons coded category membership and that the monkey imposed a category boundary on the images. But, in this type of categorization it is impossible to be certain which stimulus features are controlling categorization. Thus, it is hard to find out what the

neural activity actually represents, whether feature or category. This is a key advantage of approaching categorization in the spatial domain as it is easier to assign neural activity to the neural representation of position and category. This will help us to characterize the level of representation at which executive control is exerted.

Chapter -1

Position and Category Signals in Parietal and Prefrontal cortical networks

Introduction

Categorization refers to the brain's ability to classify different objects with similar attributes into a group. The ability to place objects into categories allows us to impart the characteristics of the category to a particular object and therefore use our knowledge about the properties of that category to guide action. For example, knowing that an object belongs to the category "knife" will instruct us to keep it away from children or not carry it into an airplane because of the two characteristics we know are common to objects that belongs to the "knife" category i.e. knives are sharp and can cut and thus, are dangerous around children and also, knives are not allowed on airplanes. Thus, categorizing the objects makes it easier to apply knowledge to perform appropriate actions and to predict the behavior of the object in various circumstances.

Categorization seems a unitary process, however there are likely to be many distinct forms, several of which have been identified based on the distinct features of cognitive tasks developed to study categorization in humans to date (Ashby and Maddox 2005). In their recent review, Ashby and Maddox divide categorization tasks into four main classes: rule-based tasks, prototype distortion tasks, information-integration tasks, and weather-prediction tasks. In prototype distortion tasks, for example, each category is defined by a prototype or exemplar object with features that define the characteristics of

that category. Subjects are initially trained to classify the prototypes, and then are challenged to classify objects with features that are intermediate between the prototypes. In this way, it is possible to discover the boundary between the prototypes that subjects have imposed to differentiate objects. In the weather prediction task, subjects are given a collection of cues that jointly instruct the probabilities associated with a binary outcome (e.g. 'rain' or 'shine'), and must combine information over the cues to make their choice or 'prediction'. Information integration tasks are the ones in which combinations of multiple features of the same object define category membership, but subjects often are not able to explicitly describe the basis for their categorical choices. Rule-based tasks are the ones that have explicit instructions of what defines the boundary between the two categories i.e., there is an explicit rule defines category membership, whereas in other types of categorization tasks, categories are often implicitly learned by trial and error and although subjects become proficient, the features that define category membership may not be explicitly known. Out of these, rule- based categorization tasks are most relevant to our experiment.

In the following section, I will describe the neural studies performed to understand categorization. The neural correlates of categorization have been studied in monkeys performing a number of different categorization tasks (Freedman, Riesenhuber et al. 2002; Freedman and Assad 2009; Crowe, Averbeck et al. 2010). In the paper by Freedman (Freedman, Riesenhuber et al. 2001) , the authors developed a prototype distortion task to investigate the neural correlates of object categorization in prefrontal

cortex. They trained monkeys to characterize computer generated 3D visual objects as cats or dogs. The images of dogs and cats were morphed together in various combinations and the monkeys used groups of distinctive features to classify each image as being a dog or a cat. Categorization as a cognitive operation is a step function of classification probability with the step aligned to the category boundary. This indicates that the subjects treat all objects within the category as equally the same, and all objects in different categories as equally different. There is a possibility that the neural signal within one category will look very alike with a sharp tuning across the category boundary and thus having lower or very different firing rate for the other category. Freedman found single neurons selective for cats, others for dogs, and found that these neurons appeared to demonstrate an all-or-none response of approximately equal strength for all members of the category. There was an abrupt change in firing rate in response to relatively small changes in stimulus properties when the stimuli crossed the boundary from dog to cat or cat to dog. The information carried by these neurons did not relate to the motor response (a lever press to indicate whether an object matched the category of the sample or not) and thus, they concluded that the activity of prefrontal cortical neurons coded category.

In object categorization tasks and similar object prototype distortion tasks, it can sometimes be difficult to determine exactly what features of the objects are determining the monkeys' categorical judgments, as objects in the two categories differ along a number of dimensions simultaneously. This in turn can make it difficult to determine

what category-selective neural activity represents – whether these neurons are selective for features of objects that differentiate categories, or the categories themselves as cognitive variables that are abstracted from stimulus features. For example, some neurons with tuning functions distributed along the feature continuum are going to exhibit category-selectivity, but they may just be part of a population of neurons coding the feature dimension on which categories are defined. Category representations could use a gain-field mechanism in the sense that neurons could retain parametric tuning for the relevant feature dimension, with the peaks of tuning functions for individual neurons distributed along that dimension but remaining fixed, whereas the amplitude of the response would be modulated by whether a stimulus belonged to a preferred category or not. Ambiguity about whether neurons code features or categories also makes it difficult to analyze the inter-relationship between signals that code features and category, to determine if feature and category-selective neurons interact during the categorization. Alternatively, neurons could code category membership as an abstract cognitive construct, independent of the underlying feature space. If features and categories are tightly correlated, it is difficult to distinguish the neural correlates of either.

In order to differentiate these alternative neural mechanisms as cleanly as possible, we developed a spatial categorization task in which we can decipher category from space. Spatial categorization tasks simplify the feature dimensions that define category membership to one or a few dimensions, these dimensions are easily identified and a good deal about how single neurons are tuned to the spatial features of stimuli is

known. A recent study by Fortes et al. took advantage of this fact and developed a spatial categorization task and recorded from parietal and prefrontal cortex during task performance (Merchant, Crowe et al. 2011). The authors developed a prototype distortion categorization task based on the position of a visual stimulus, and trained monkeys to classify a horizontal bar presented at different vertical positions within a box as either 'high' or 'low'. Monkeys were trained to classify high and low, a horizontal bar of light was presented at either high or low positions in a video display as a sample stimulus and the choice was made by pressing when the choice matched the vertical position of the sample and then on a minority of trials, monkeys were probed with bars of intermediate height. These authors found that when only trained with prototype stimuli, monkeys inferred the position of the boundary dividing high and low categories, and were able to categorize bars of intermediate height spontaneously. If the monkey chose correctly, it was rewarded with a drop of juice. Further, this study demonstrated that single neurons in prefrontal and also posterior parietal cortex exhibited category-selective patterns of activity, with different neurons preferring 'high' and 'low', and showing the step-like change in firing rate as bars moved from one category to another that was consistent with the pattern of behavioral report of the monkey. Comparing the pattern of activity observed between these areas, the study reported that the spatial category signal appeared much earlier in PFC. The results of this study provide evidence that prefrontal cortex provides an early top-down spatial category signal, and that this signal does not emerge in posterior parietal area 7a until the monkey has to express that categorical decision through a response. Whereas opposite dynamics was reported by Freedman and

colleagues, indicating the parietal cortex is originator of category signal and the category signal is stronger and earlier in area LIP (in parietal cortex) in comparison to prefrontal cortex (Palmer, McDonough et al. 2012). Also, to dissociate our study from the recent paper by Swaminathan et al., ours uses spatial relationships, and more importantly, working memory. Specifically, category is determined, in part, on the basis of information stored in working memory (the position of the categorized stimulus), and this could better dissociate category processing from sensory processing in a way that could influence network dynamics.

In the Freedman and Assad study (Freedman and Assad 2009), single neurons were studied when there was a one-one mapping between directions and categories. Neurons exhibited step-like tuning to direction, but it was not possible to determine unambiguously whether that activity reflected tuning to motion direction as a stimulus feature or motion category abstracted from direction. There are therefore still unanswered questions outstanding regarding the network dynamics that generate category-selective activity in the parietal-prefrontal network and the degree to which position and category information can be fully parsed to best reveal the underlying neural mechanisms.

To better make that distinction, we developed a novel spatial categorization task in which monkeys were required to assign the same physical stimulus location to opposite spatial categories over trials, enabling us to clearly dissociate stimulus position

and stimulus category as behavioral variables. In this task, spatial categories are defined on the basis of a spatial relationship between two sequentially presented stimuli, a dot and a line. Monkeys have to determine the spatial category of the dot based on the spatial relationship to the line. This represents an extension of ongoing work based on the neural coding of spatial relationships in posterior parietal cortex (Crowe, Chafee et al. 2004; Crowe, Averbeck et al. 2005; Chafee, Averbeck et al. 2007). Unlike the study by Fortes, Georgoupolos and colleagues, in our study we explicitly cue the boundary dividing left and right spatial categories by presenting a vertical line as a cue to indicate the location of the category boundary. This allows us to both force the monkey to assign the same position in the visual display to alternative categories (by shifting the position of the line) and sets the stage for the study of the rule-dependent categorization of space (in Chapter 2).

In this experiment, we will address the question of how categories are represented at cellular level. Spatial feature i.e. position is a visual characteristic of the visual stimuli (an intrinsic property), whereas the category of the stimulus is an abstract property assigned to the stimulus by a cognitive process as a function of context. This is the key advantage of approaching categorization in the spatial domain as it is easier to separate features and category of the stimuli, which in turn can help us to isolate the neural signal representing each.

By virtue of the design of our task 1) spatial position and spatial category will act as independent, dissociated variables, 2) we will be able to characterize the activity of individual neurons as the boundary value changes, and therefore, 3) we will be able to characterize how neural activity changes to reflect the alternate assignment of one stimulus position to opposite spatial categories to more fully dissociate neural signals coding in this rule-based task. This will help us determine the degree to which spatial category information is abstracted from spatial feature information at the single neuron level.

Materials and Method

Animals Two male rhesus macaque monkeys (5-8kg) were trained. Neural recordings were performed simultaneously from prefrontal and posterior parietal cortex. All surgical and experimental procedures associated with surgery and neurophysiological recording conformed to NIH Guidelines and protocols approved by Internal Animal Care and Use Committee (IACUC) at the VA Medical Center and the University of Minnesota. Gas anesthesia (Isoflurane, 1-2%) was used in an aseptic surgical procedure in which we placed a craniotomy centered over the principal sulcus (Brodmann's area 46) in the left prefrontal cortex, and another craniotomy centered over the posterior aspect of the inferior parietal gyrus (Brodmann's area 7a) in the left posterior parietal cortex, of both monkeys (Fig. 3 d, e). We placed titanium screws in the skull, and attached recording chambers (either 7 or 13 mm i.d.) and posts for head fixation to the screws using surgical

bone cement. We administered post-surgical analgesia for several days (Buprenex, 0.05 mg / kg Bid, i.m.).

Task Dynamic Spatial Categorization – DYSC-Shift Figure 3 illustrates the DYSC- shift task. The trial started with a fixation dot (0.25 degrees white spot), which the monkey was supposed to fixate during the whole trial. If at any point, the gaze of the monkey deviated more than 1.5-2.5 degrees from the fixation dot, the trial was terminated without a reward and a tone was presented to indicate an error. After 500 ms of central fixation, a sample stimulus (a 0.5 degrees white spot) was presented for 400ms. The sample was presented at a location selected from a rectangular array of 18 positions (2 rows by 9 columns). The spacing between adjacent sample locations in the grid was 2.5° in the horizontal dimension and 5° in the vertical dimension. We presented a vertical line, serving as a category boundary, that crossed the grid of sample locations either 5° to the left or 5° to the right of the gaze fixation target (Fig. 3c). On trials in which we presented the boundary to the left of the fixation target, we selected the sample from the 12 positions in the grid that were immediately to the left and right of the boundary location. In this case, the sample locations to the right of the boundary made up the central 6 positions in the grid. On trials in which we presented the boundary to the right of the fixation target, we again selected the sample from the 12 positions in the grid immediately to the left and right of the boundary, but in this case, the central 6 positions in the grid now fell to the left of the boundary. In this way, the central 6 sample positions located in the middle of the sample grid could either fall to the relative left or right of the

category boundary over trials. This ensured that the brain would be required to assign samples appearing at fixed retinal locations in between the two boundaries alternatively to *LEFT* and *RIGHT* spatial categories, across trials. Following the sample was a 900 ms delay, after which we presented the boundary cue (a vertical line) for 400 ms. The position of the vertical boundary cue instructed which portions (or more specifically which horizontal positions) of the sample grid would be defined as LEFT and RIGHT based on the position of the sample relative to the boundary. For the majority of the trials we used 2 boundary positions (± 5° degrees either to the left or right of the central target). In some trials we used 3 boundary positions (± 10° degrees to left and right of central target and one right at the central location). After the boundary cue disappeared, there was again a delay for 900ms, after which two choice stimuli (small white circles identical to the sample) were presented in sequence. The two choice stimuli fell to either side of the category boundary, one that matched correctly the spatial category of the sample stimulus (and was located on the same side of the boundary) and the other that was located in the opposite spatial category relative to the sample and on the opposite side of the boundary (the nonmatch). The order of match and nonmatch choice stimuli was randomized over trials. The monkeys were rewarded if they pressed a response key during the period of time that that matching choice stimulus was visible. If the first choice was the match (1/2 of trials at random), monkeys were rewarded for pressing the response key during the first choice period, in which case the trial terminated (Fig. 3b). If the first choice was the nonmatch, the monkey had to withhold responding during the 1st choice period, and depress the response key during the 2nd choice period while the

match was visible (Fig. 3a). The response key was a pedal the monkey depressed with its right foot (movement direction did not vary over trials, and response timing was undetermined before presentation of the choice stimuli).

In the DYSC-shift experiment, we presented first the sample followed after a delay by the boundary cue. Thus, a key feature of the DYSC shift task was that the monkey had to combine information stored in spatial working memory (the sample position) with information provided by visual input (the boundary cue) in order to compute the spatial category of the sample. This helped to dissociate neural signals coding the spatial category of the sample (evoked presumably by the appearance of the boundary cue) from neural signals coding the position of the sample (evoked earlier in the trial by the presentation of the sample). The fact that the same stimulus locations in the central portion of the grid could fall alternatively to the left or right of the boundary cue (depending on its position), helped further to dissociate signals coding spatial position and spatial category.

Neural Recording We recorded neural activity in prefrontal and posterior parietal cortex simultaneously using two 16 electrode Eckhorn microdrives (Thomas Recording, GMbH; Giessen, Germany). Electrodes were 70 μm o.d. glass coated platinum iridium fibers (impedance 1-2 MΩ). Each electrode was advanced independently under computer control. The electrical signal from each electrode was amplified (5000X) and filtered (between 0.5 Hz and 5 KHz). The waveforms of the action potentials of individual

neurons were isolated on-line by two people using a combination of time-amplitude window discriminators (Bak electronics, Mount Airy, MD) and waveform discriminators (Alpha Omega Engineering, Nazareth, Israel). The timing of spike occurrence was stored to disk with 40 μs resolution (DAP 5200a Data Acquisition Processor, Microstar Laboratories, Bellevue, WA). We typically isolated the action potentials of 15-25 neurons simultaneously in each cortical area.

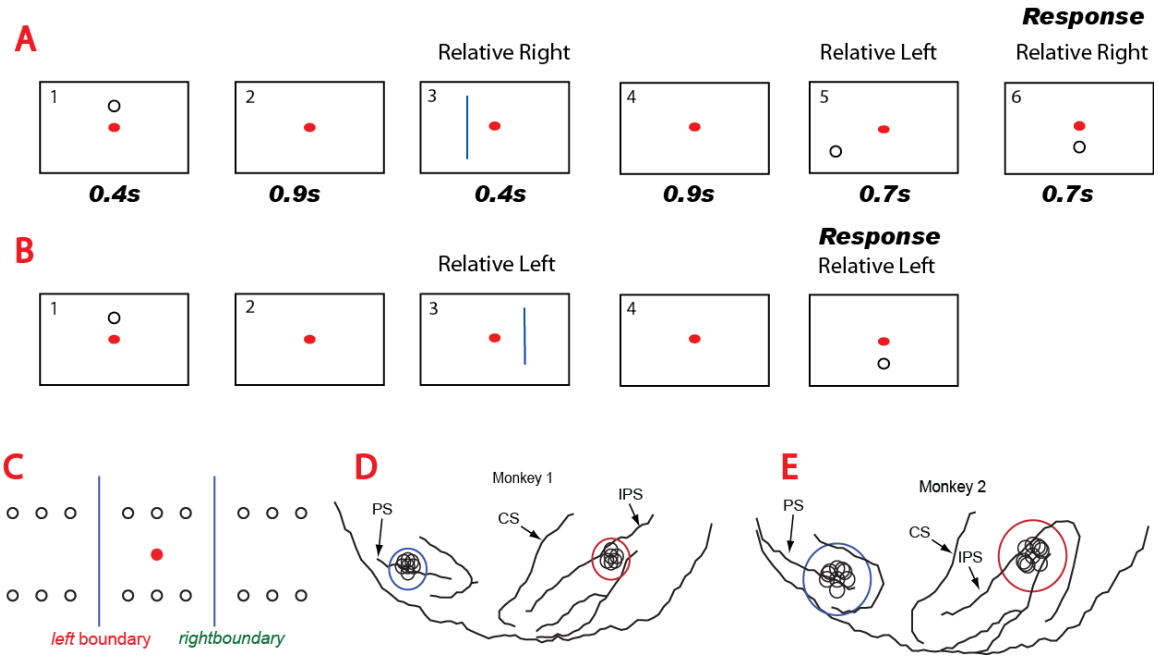

Figure 3. DYSC-Shift task design, stimulus geometry and recording locations. ***A.*** RIGHT spatial category trial. Trial starts with gaze fixation of a central target (red dot, not shown) for 0.5 seconds. After the fixation period, a sample stimulus was presented for 0.4 s (open circle; panel 1). After the sample period, a 0.9s delay period followed (panel 2). After the delay, we presented the boundary cue (blue line panel 3). On this trial, the position of the remembered sample was located to the relative right of the boundary cue, placing the sample in the RIGHT spatial category. After a second 900 ms delay (panel 4), two choice stimuli were presented, one to the left (panel 5) and one to the right (panel 6) of the boundary cue, in random sequence. The correct response on this trial was to press the response key when the visible choice was located within the same spatial category as the sample (panel 6 in this example as the choice, like the sample, was located in the category RIGHT relative to the boundary). ***B***. LEFT spatial category trial. A sample presented at the same location (compare panel 1 in A and B) is now assigned to a different spatial category (relative left), because a change in boundary position has

altered the spatial relationship between sample and boundary. ***C***. Stimulus array. Sample and match stimuli appear in a 2 by 9 rectangular grid of positions with 2.5° spacing between rows and 5° spacing between columns. There are two boundary locations at -5° and +5° degrees relative to the central fixation target. ***D, E.*** Locations of neural recordings in parietal and prefrontal cortex of Monkeys 1 and 2 relative to positions of the principal sulcus ('PS'), central sulcus ('CS') and intraparietal sulcus ('IPS') as reconstructed from structural MRI images. The perspective is a top-down view of the left cerebral hemisphere. Smaller circles (black outline) within the chambers (colored outline) indicate the approximate location of the circular tip of the multi electrode matrix for a single experiment.

Data Analysis

ANOVA-based definition of position and category neurons. To identify neurons with activity that varied as a function of the position of the sample stimulus, we measured firing rate during the sample and following delay 1 periods before the onset of the boundary cue (Fig. 3a, b; panels 1 and 2), and evaluated the significance of the influence of sample horizontal position on activity in this interval (one-way ANOVA with sample horizontal position with 6 levels as the single factor). To identify neurons with activity that varied as a function of the spatial category of the sample (LEFT or RIGHT), we measured firing rate during the boundary cue and subsequent delay 2 periods (Fig. 3a, b; panels 3 and 4), and performed a 2-way ANOVA (Type III sum of squares) on this activity in which the factors were spatial category and boundary position. Because we presented the sample first and then the boundary cue in these experiments, the category of the sample (LEFT vs. RIGHT) was only defined after the boundary cue appeared. In the above analysis of category-selective activity, we used the firing rate during the

sample presentation and first delay as the covariate - so as to remove any carry-over effect on firing rate due to sample position. Neurons were defined as category selective when their activity was significantly influenced by spatial category not by the position of the boundary cue or the interaction between boundary and category. This definition of category-selective neurons would identify neurons in which the difference in activity between LEFT and RIGHT spatial categories did not vary as a function of the absolute position of the boundary cue or the sample stimulus, but was based rather on the position of the sample relative to the boundary. Neurons meeting these criteria would code categories in a way that exhibited translation invariance.

Spatial tuning of position neurons To visualize the spatial tuning of individual position neurons, we selected neurons with activity varying significantly ($p < 0.05$) as a function of sample horizontal position in the above ANOVA, and plotted firing rate over the 18 sample positions in the grid as a 3D surface with height representing firing rate, and interpolating values of the surface between sample positions. To construct an average population spatial tuning function of position neurons, we selected neurons with activity varying significantly as a function of horizontal sample position ($p < 0.05$), normalized the spatial tuning of each neuron to its peak, aligned the tuning functions across neurons to the preferred horizontal position of each neuron, and then averaged the functions over all neurons in the population.

Time course of position and category signals. To illustrate the time course of the neural signals coding sample position and spatial category, we constructed normalized population spike density functions (SDF) of their activity. We selected neurons with activity relating significantly ($p < 0.05$) to the horizontal position the sample, the horizontal position of the boundary, or the spatial category of the sample (by the ANOVA criteria above), determined the preferred sample position, boundary position, or spatial category for each neuron (associated with the peak firing rate), and then constructed separate spike density functions using activity recorded on preferred and nonpreferred trials for each neuron by convolving the spike train on each trial with a Gaussian kernel of width $\sigma = 20$ ms. We then averaged the single-trial SDF over preferred and nonpreferred trials for each neuron, and normalized both average SDF to the peak rate on preferred trials for that neuron. Finally, we averaged the normalized single neuron SDF across all neurons in the population to produce the average population SDF. We generated separate average population SDFs for position neurons using different populations selected on the basis of whether activity related significantly to sample position during the sample period, the delay period 1 following, or the boundary cue period.

Correlation in baseline activity The spatial categorization required by the DYSC-shift task presumably requires that neurons which code position interact with neurons that code the categories to which those positions have been assigned. In order to determine

whether position and category neurons might interact synaptically, we computed the correlation coefficient using the activity of pairs of neurons measured in the inter trial interval. We selected position and category neurons on the basis of the ANOVA criteria described above. We focused our analysis on baseline activity in order to evaluate noise correlation in pairs of cells as an indirect measure of the presence of synaptic interaction between neurons (correlated fluctuation in activity during the inter-trial interval is correlated noise in the sense that it occurs before stimuli were presented, and neurons were driven by sensory input to carry signals related to the task). We wanted by this analysis to determine if there is a connection/interaction between position and category neurons intrinsically, as our hypothesis was that these two types of cells need to be connected in order for them to communicate to transform a position signal into a category signal during the task. To ensure that measures of correlated activity reflected true interactions between neurons, we compared the number of significantly correlated neuron pairs obtained when the two neurons were recorded simultaneously or at different times. Finally, to determine whether correlated baseline activity might relate functionally to the signals that neurons carried during task performance, we compared the mean correlation coefficient and the number of significantly correlated neuron pairs when the horizontal position preferred by one neuron belonged within the horizontal spatial category preferred by the other neuron, versus the converse case in which the position preferred by one neuron fell outside the category preferred by the other.

BMI (Boundary Modulation Index) To quantify the degree to which the neural signal coding spatial category was invariant with respect to the absolute spatial locations of the boundary and the sample stimuli (e.g. the degree to which category-selective activity was of comparable strength when the boundary was left and right) we calculated a coefficient we termed the Boundary Modulation index (BMI). We first divided trials according to the position of the category boundary (left and right), and computed the difference in firing rate between preferred and nonpreferred spatial categories for each boundary position. We then computed the BMI for each neuron as (A – B) / (A + B), where A and B are the category signals (difference in rate between preferred and nonpreferred categories) observed at each boundary location. We further defined the A term above to be the category signal observed when the boundary was on the same side as the category preference for each neuron (for example the left boundary for a neuron with a left spatial category preference). The values of BMI ranged from +1 to -1; +1 meaning that the neuron was only active for the congruent boundary (on the same side as its category preference), and -1 indicating that the neuron was only active for the incongruent boundary. A BMI value of 0 indicated that there was no difference in the category-selective activity of the neuron between the two boundaries, which in this case indicates a category signal that exhibits translation invariance and is equally strong regardless of the absolute position of the samples and the boundary in retinal coordinates.

Decoding Position and Category We also applied a pattern classification analysis (Klecka 1980; Johnson and Wichern 1998; Crowe, Averbeck et al. 2010) to decode

sample position, boundary position, and spatial category on each individual trial of the DYSC task from the pattern of population activity in parietal or prefrontal cortex observed on that trial. The proportion of trials for which the analysis identified the correct values for the spatial variables of sample horizontal position, boundary horizontal position, and spatial category, given the stimuli displayed each trial, provided a measure of the strength with which population activity represented each variable in each cortical area. We selected neurons for inclusion in the populations used in the decoding analysis based on the significance of the influence of the decoded variable on firing rate. We first applied the analyses of variance and covariance to neural activity as described above and then ranked neurons in both cortical areas according to the significance (p-value) of each factor in the analysis. In order to decode sample position, for example, we selected the 20 neurons in parietal and prefrontal cortex of each monkey most significantly influenced by sample position (having the lowest p-values) in the ANOVA (decoding position from population activity patterns in the two monkeys separately and then averaging the results). Similarly, to decode boundary position and spatial category, we selected the 20 neurons most significantly influenced by these variables in each cortical area in each monkey to perform the decoding analysis. This approach held the number of neurons included in the analysis in parietal and prefrontal cortex constant (removing the number of neurons as a factor influencing decoding accuracy). It also maximized decoding accuracy in both cortical areas by focusing on the neurons with the strongest signals in each area. We performed the decoding analysis in a time-resolved manner, using firing rates measured in 50 ms time bins throughout the trial. We represented the pattern of

population activity in each time bin as a population activity vector consisting of a list of the firing rates of each neuron in that bin. We then applied 5-fold cross validation to decode the variable of interest from population activity. On each of 5 iterations, we used 4/5 of trials as the training data to train the classifier, and then used the resulting classifier to decode the variable of interest from population activity patterns on the remaining 1/5 of trials. This helped to ensure that the classification was robust and that comparable levels of decoding accuracy could be expected if the classifier were applied to new neural data. To train the classifier using the training trials, we computed the mean population activity vector associated with each level of the decoded variable as well as the within-group covariance matrix between neurons pooled across levels of the decoded variable. We then modeled the patterns of activity coding each category as a multivariate normal distribution of population activity vectors, with mean equal to the mean vector observed for a given level of the decoded variable, and variance quantified by the pooled covariance matrix (Averbeck, Crowe et al. 2003). We then decoded the spatial variable of interest from population activity on each test trial by giving that variable the value associated with the highest posterior probability in the analysis. We computed the posterior probabilities using the *classify* function in the Matlab statistical toolbox (The Mathworks Inc., Natick MA). At each time step in the decoding analysis, we represented the activity of each neuron by three successively measured firing rates: the mean firing rate of each neuron in the current 50 ms time bin, and in the two preceding time bins. We then concatenated the resulting three successive firing rate measurements for each neuron over all neurons in the population to construct the population activity vector at each time

step and performed the decoding analysis using these population activity vectors. (Each population vector constructed in this manner reflected short-term fluctuations in the firing rate of each neuron over a short interval at each step of the decoding analysis and generally improved decoding accuracy). We then plotted the proportion of trials in which the analysis returned the correct value of the decoded variable as a way to measure the strength of the population signal coding that variable at each time step, advanced the window 50 ms, and repeated the decoding analysis. Connecting the points generated by this stepping procedure produced a decoding accuracy time course that measured fluctuation in the strength of population signals coding sample position, boundary position, and spatial category over time. This provided a basis to compare the timing of these signals in parietal and prefrontal cortex.

Results

Behavioral Performance

Figure 4 plots the proportion of trials that each of the two monkeys classified the sample stimulus as belonging to the spatial categories RIGHT (Fig. 4; blue bars) or LEFT (Fig. 4; red bars) as a function of the horizontal position of the sample, and the horizontal position of the boundary cue. In the DYSC-shift task, we had two different horizontal locations for the boundary which required that monkeys reclassify the same region of space (between the two boundaries) into LEFT or RIGHT spatial categories, and this is important as it helps to dissociate visual neural signals coding the retinal position of the sample stimulus from others coding the category the brain places the sample into.

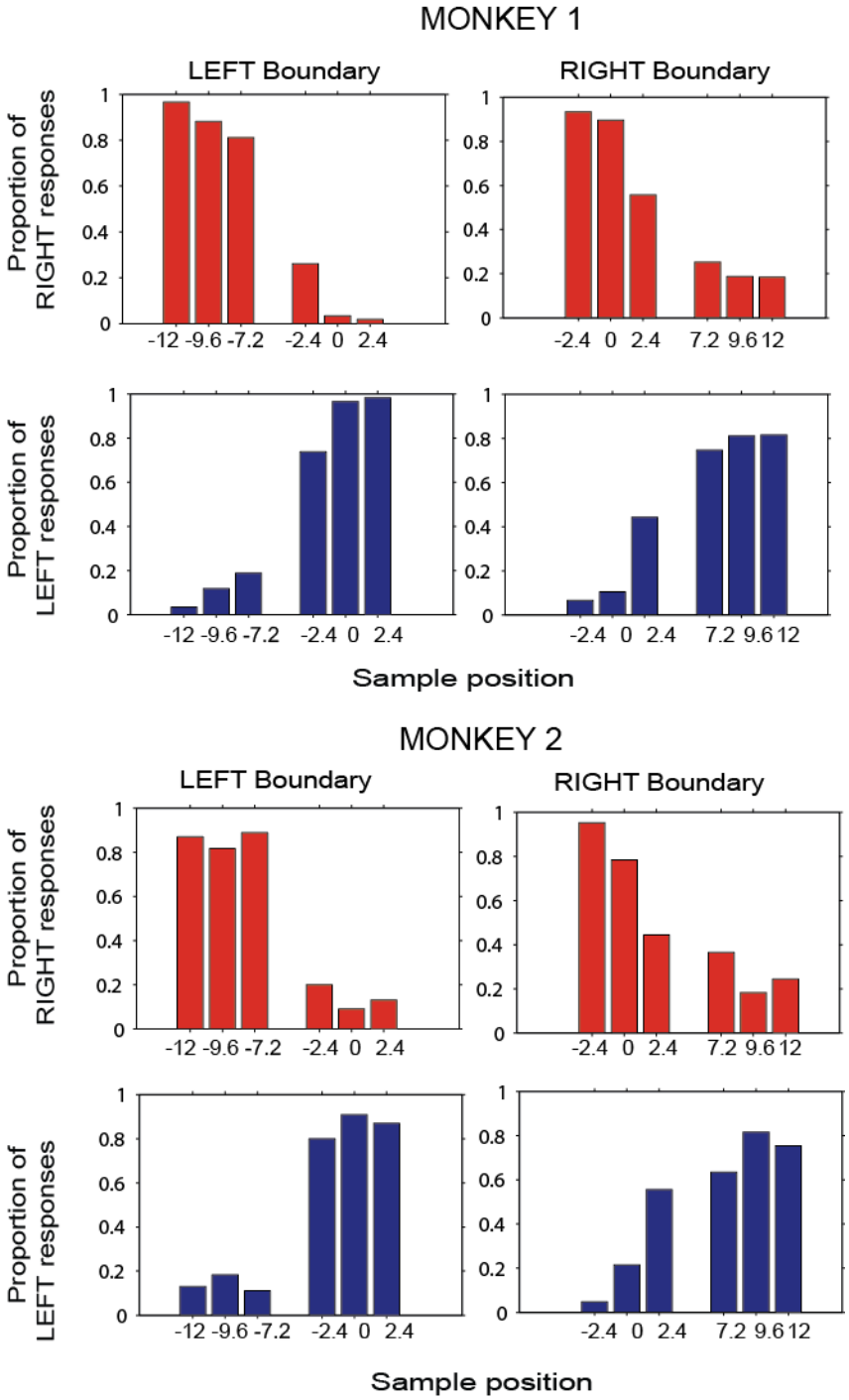

Figure 4. Behavioral performance of Monkey 1 (upper four panels) and Monkey 2 (lower four panels) on the DYSC-shift task . Histrograms illustrate the percentage of trials that monkeys classified sample stimuli as belonging to the LEFT (red bars) and RIGHT (blue bars) spatial categories as a function of the horizontal position of the sample (along the

abscissa in each panel), and the horizontal position of the boundary (left and right columns in this figure illustrate behavioral data on left and right boundary trials, respectively).

Generally, monkeys selected the choice stimulus matching the category of the sample (located to the same side of the boundary line) with high probability, and there was a relatively abrupt reduction in response probability as the sample stimulus crossed to the other side of the category boundary. However, this was an exceedingly difficult task to train, most likely because monkeys had to compare a location stored in spatial working memory (the position of the sample stimulus) with a visible cue (the category boundary) which given the limited area of the stimulus array and the close proximity of sample and boundary positions in the grid, proved to be a a difficult spatial judgment. Consequently, categorization performance suffered somewhat, and in some cases, monkeys exhibited a systematic bias. For example, when the boundary was at the right position (Fig. 4; right column), both Monkey 1 and 2 more frequently misreported samples just to the left of the boundary as belonging to the incorrect RIGHT category. This could reflect the difficulty of classifying one stimulus as being located to the relative left or right of another when one of them is stored in spatial working memory. Monkey 1 performed the DYSC-shift task at 80% while Monkey 2 performed at 75% accuracy. Note also, in spite of these inaccuracies, the behavioral data provide clear evidence that the monkeys shifted the position of their perceptual dividing line between categories in response to the change in the position of the boundary cue.

Neural database

We recorded the activity of a total of 758 neurons, 381 neurons in parietal cortex and 377 neurons in prefrontal cortex of two monkeys as they performed the DYSC-shift task in which we presented the sample stimulus first followed by the boundary and in which we randomized the position of the boundary over trials. In Monkey 1, we recorded the activity of 135 parietal neurons and 148 prefrontal neurons during DYSC-shift. The corresponding numbers in Monkey 2 were 246 parietal and 229 prefrontal neurons respectively.

Distribution of neurons coding position, rule and category

To identify neurons in which activity related significantly to sample position, boundary position, and spatial category, we performed a series of analyses of variance and covariance. We defined position neurons as those in which firing rate during the stimulus 1 and delay 1 periods varied significantly (ANOVA; $p < 0.05$) as a function of the horizontal position of the sample stimulus. We defined boundary neurons as those in which firing rate during the stimulus 2 and delay 2 periods varied significantly as a function of the horizontal position of the boundary (ANOVA; $p < 0.05$). We defined category neurons as those in which firing rate during stimulus 2 and delay 2 periods varied significantly (ANCOVA; $p < 0.05$) as a function of the horizontal category of the sample stimulus (left or right) One objective in this analysis was to evaluate category signals after removing the influence of sample position on firing rate. For that reason, we included firing rate during stimulus 1 and delay 1 periods as a covariate in this analysis.

The rationale for this was that it would factor out variation in firing rate due to sample position that would emerge in the stimulus 1 and delay 1 periods, and consider the influence of category independently of position effects.

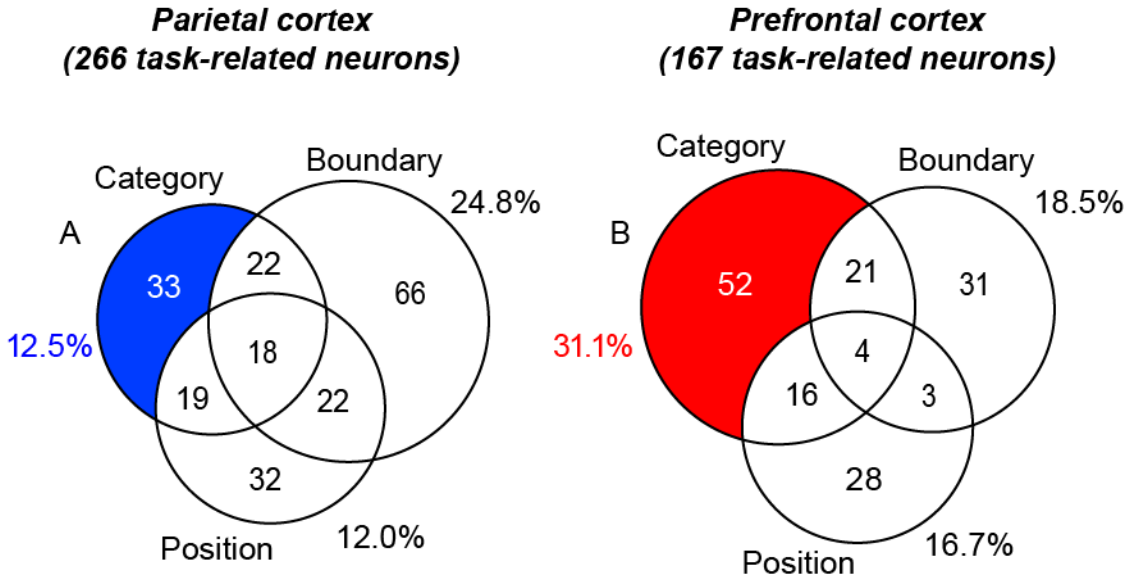

Figure 5. Venn diagram showing number of neurons significant for spatial category (left or right), sample position and boundary position (left or right) in the ANOVA/ANCOVA. ***A.*** Distribution of neurons with activity related to each factor in parietal cortex. ***B.*** Distribution of neurons with activity related to each factor in prefrontal cortex.

In Figure 5, the numbers of neurons with activity that relates to the horizontal position of the sample, boundary position and spatial category are plotted in two separate Venn diagrams, one each for parietal and prefrontal cortex. We found that there were neurons coding all three behavioral variables in both of the cortical areas which indicates that the processing involved in this task is thoroughly distributed. However, we found that more than double the proportion of task-related neurons were significantly related to category but not the other factors in prefrontal cortex in comparison to parietal cortex (Fig. 5, colored regions). This difference in pure category neurons between parietal and

prefrontal cortex was significant (chi-square test, p = 0.0294). When considering all category neurons, including those also related to boundary or sample position, there were still proportionally more category neurons in prefrontal cortex (55.7%) than parietal cortex (34.6%), but this difference was no longer significant (chi-square test, p=0.86). Conversely, we found a higher proportion of task related neurons significantly related to boundary position in parietal (24.8%) in comparison to prefrontal cortex (18.5%;chi-square test, p=1.0012e-008). Similarly the difference between the proportion of neurons exhibiting a position signal was also significantly different between the two cortical areas (chi-square test, p=0.00025). The proportionally larger number of neurons reflecting the position of the sample in prefrontal cortex may reflect the requirement to store this position in spatial working memory.

Comparison of neural signals coding sample position in parietal and prefrontal cortex

We observed single neurons with activity coding the spatial position of the sample stimulus both in parietal and prefrontal cortex. For example, the neuron illustrated in Fig 6b and c in parietal cortex recorded during the DYSC-shift task was activated during the sample period by the appearance of the sample stimulus in the left positions of the array (Fig 6b and c; positions -12.0° to -2.4°). We plotted one raster for each sample x-position, collapsing across the two y-positions at each x. The firing rate of the neuron is substantially increased during the sample stimulus period, demonstrating the visual sensitivity of this parietal neuron.

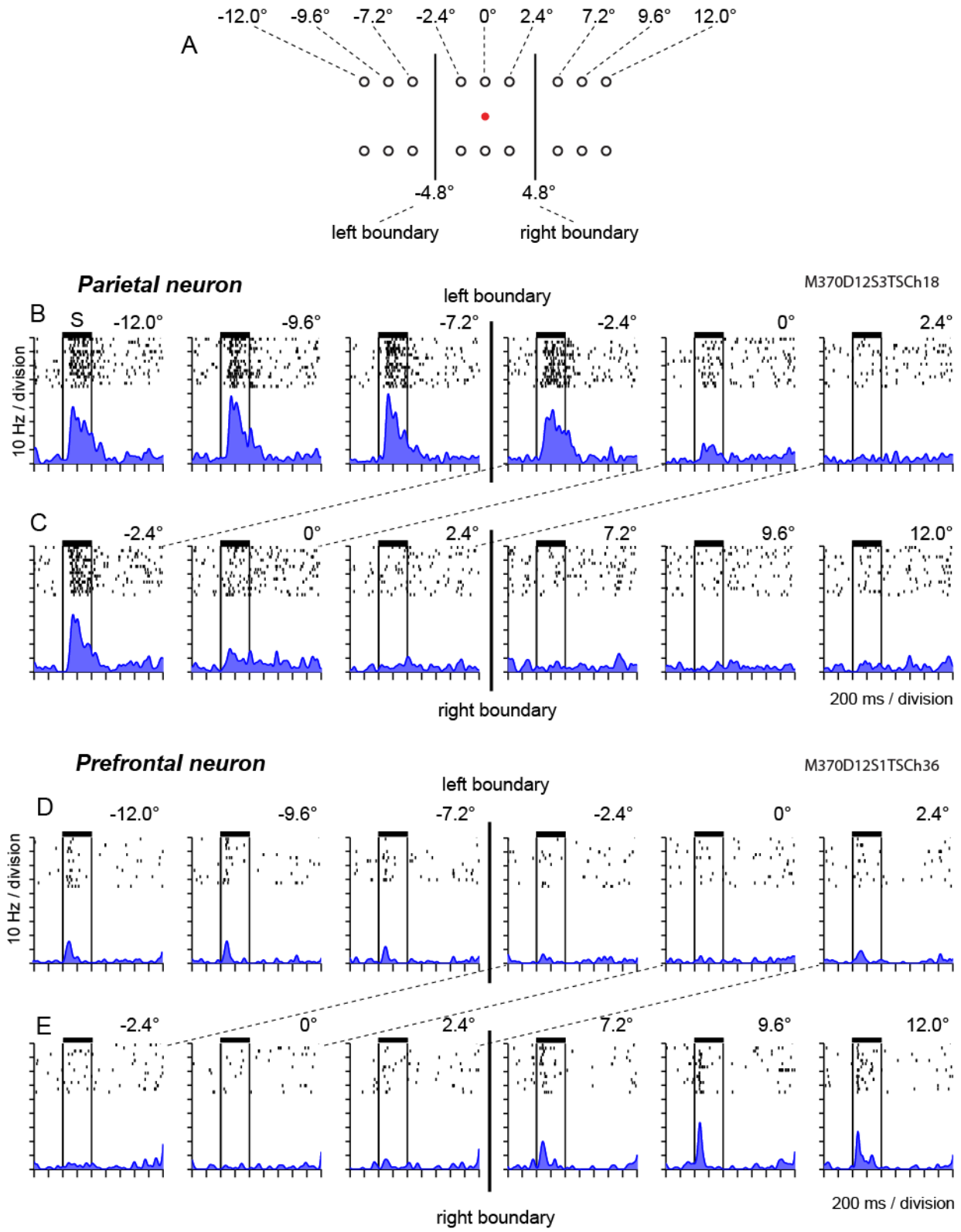

Figure 6 – Rasters of neural activity coding sample stimulus horizontal position. ***A*** Schematic of the stimulus array employed showing the relative locations of boundary

cues (vertical lines) and sample stimuli (open circles). Samples were presented in a rectangular grid of 9 horizontal by 2 vertical positions. ***B, C.*** Activity of a parietal neuron that was activated by presentation of the sample at left spatial positions in the array. ***B.*** Activity of this neuron on left boundary trials. ***C.*** Activity on right boundary trials. ***D, E.*** Activity of a prefrontal neuron that was selectively activated by presentation of the sample at right spatial positions in the array. ***D.*** Activity of this neuron on left boundary trials. ***E.*** Activity on right boundary trials.

Similarly, the prefrontal neuron shown is activated by the presentation of the sample stimulus in the right positions in the array (Fig 6d and e; positions 7.2° to 12.0°), although the response in much weaker in this prefrontal neuron.

Neural signals reflecting the position of the sample were found in multiple task epochs and were not confined to the sample period. For example, we found that there were populations of neurons in both parietal (Fig. 7, red) and prefrontal cortex (Fig. 7, blue) in which activity varied as a function of sample position during different trial epochs. For example, neurons in Fig. 7a were activated by the presentation of sample stimulus during the sample period, while neurons in Fig 7b were activated during the delay-1 period and neurons in Fig 7c during stim-2 presentation (when the boundary was shown). The important thing to note here is that the sample was not visible during delay-1 and stim2 periods, and we removed the cells whose activity varied as a function of the boundary position and category from these populations (Fig. 7c). This shows that the position signal was carried across time from stim-1 to stim-2 periods, perhaps because it is needed to compute the category of the sample along with the location of the boundary. Our hypothesis is that the position signal stored in working memory by prefrontal and parietal neurons is combined with a boundary signal to produce category information.

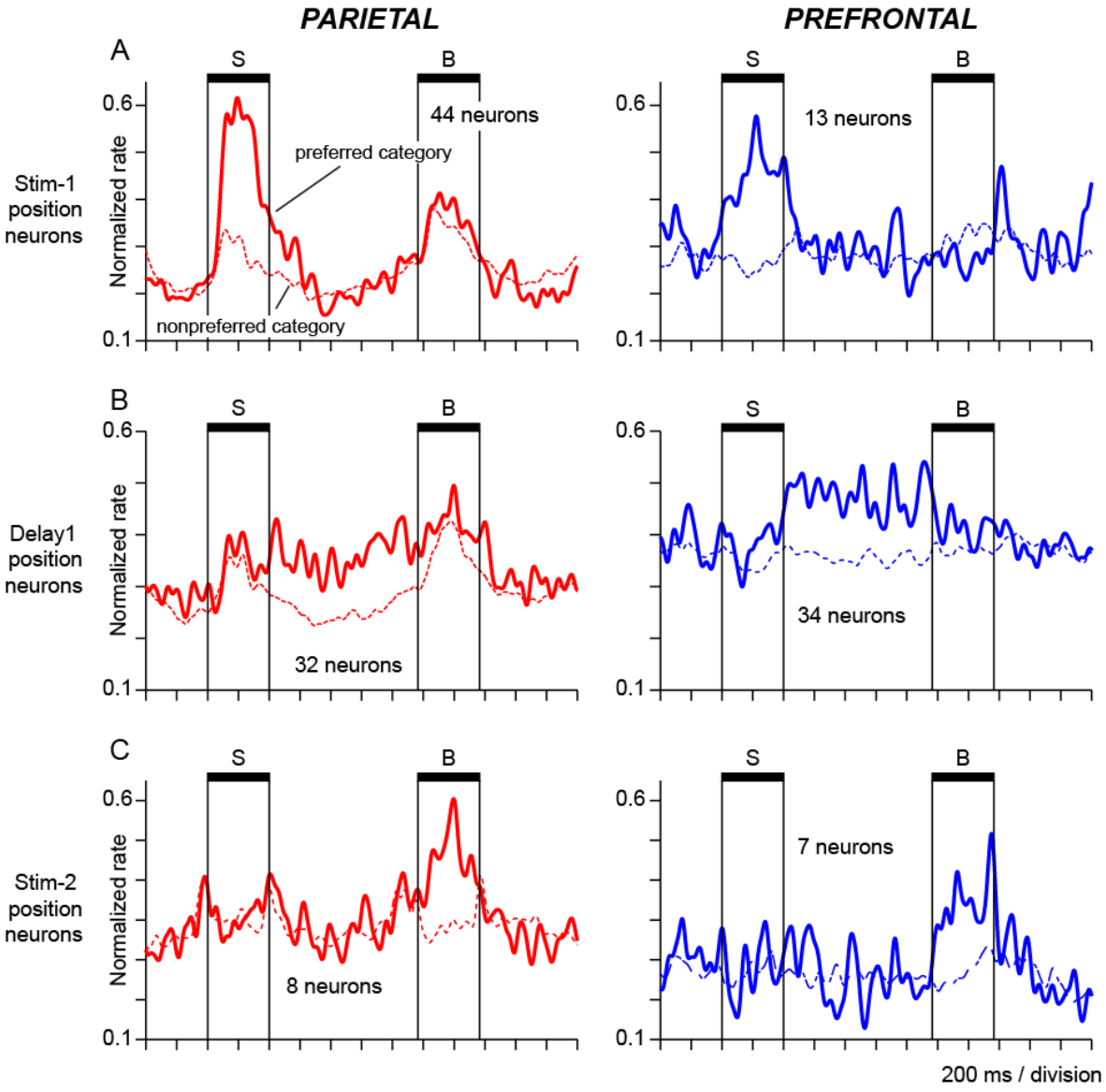

Figure 7. Normalized population spike density functions of position neurons in parietal cortex (red) and prefrontal cortex (blue). Cells were included in the populations if their activity varied significantly ($P < 0.05$) with the horizontal position of the sample stimulus during the ***A)*** stim-1 period, ***B)*** delay-1 period, ***C)*** stim-2 period.

To visualize the spatial tuning properties of position neurons, we plotted their firing rate for different sample horizontal and vertical positions in the array and interpolated the firing rate between positions to create a surface plot (Fig. 8). Fig 8a and b represent two neurons that were active in the stim-1 period whereas Fig. 8c and d

represent two neurons that were active in the stim-2 period (these neurons were not significantly influenced by boundary position or spatial category factors). These neurons were spatially tuned for essentially one position of the sample array – firing rate falls off as the distance from this preferred position increases. Thus, these cells were sharply tuned for position.

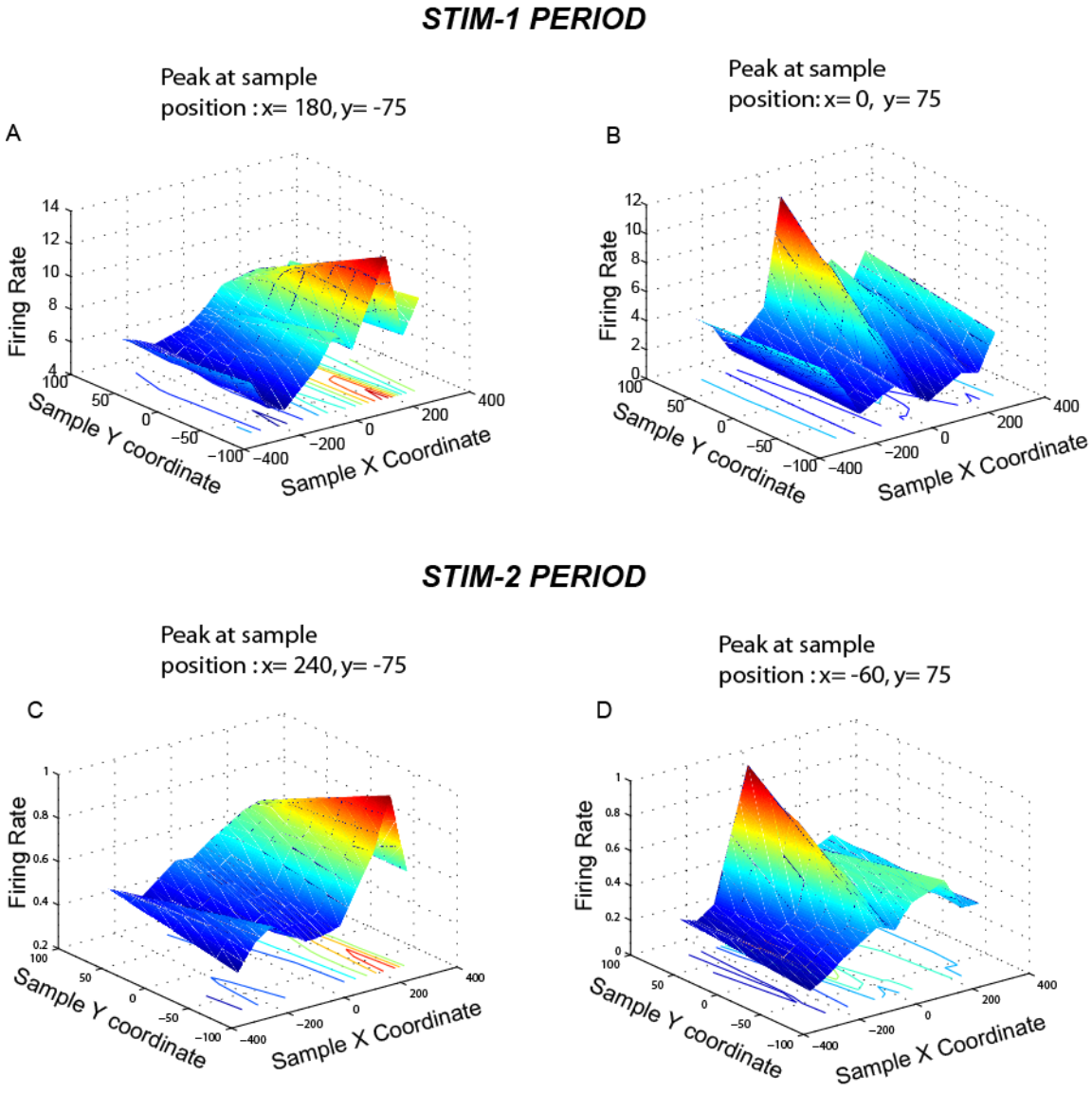

Figure 8 Spatial Tuning plots for position neurons active in the stim-1 (***A, B***) and stim-2 periods (***C, D***). The firing rate of each neuron was significantly influenced by sample position (ANOVA, $p<0.05$). The neurons in the stim-2 period were not significantly influenced by boundary position or spatial category. (Sample position is expressed in pixel coordinates in the visual display, ± 300 pixels corresponds to ± 12° of visual angle, see stimulus array in Fig. 4a).

To evaluate the spatial tuning characteristics of position neurons at the population level, we constructed normalized population spatial tuning functions by aligning the normalized spatial-tuning functions of each neuron to the preferred horizontal position of each neuron and then averaging them (Fig. 9a). We evaluated spatial tuning with respect to horizontal position because this is the stimulus dimension that we were able to dissociate from spatial category in the DYSC-shift task by shifting the horizontal position of the category boundary. The population spatial tuning functions of prefrontal neurons (Fig. 9a; blue) are very similar to parietal neurons (Fig. 9a; red). In both areas we found that the spatial tuning curve was fairly sharp.

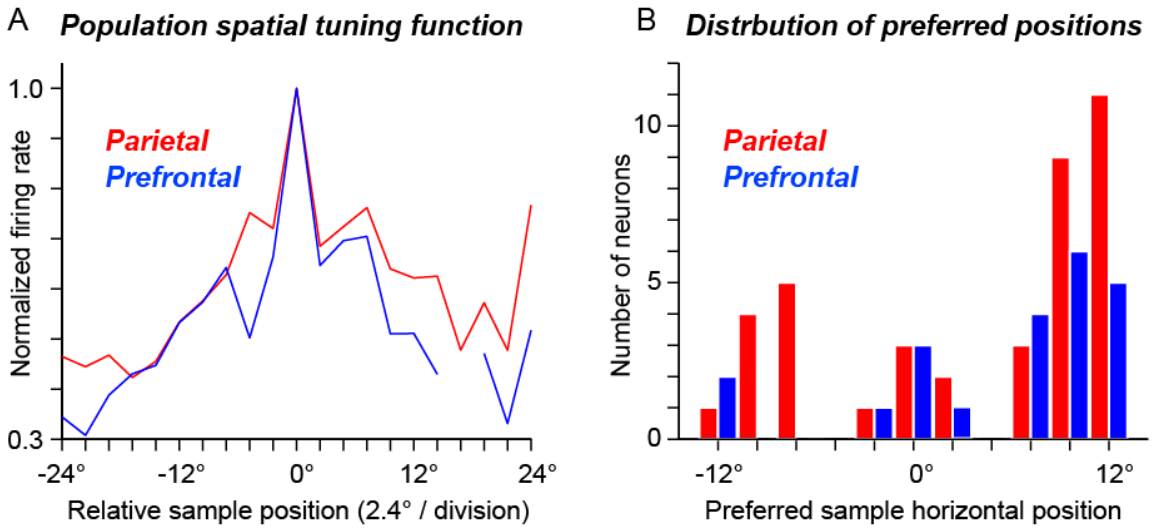

Figure 9. Population tuning curves and distribution of preferred positions for all position neurons in parietal (red) and prefrontal (blue) cortex. ***A.*** Normalized average population spatial tuning functions aligned to the peak response of each neuron at the 0° location.

B. Frequency distribution of neurons with peak responses to the sample stimulus at various horizontal positions in the array.

We calculated the total number of position neurons preferring each horizontal position in the sample array for both parietal (Fig. 9b; red) and prefrontal (Fig. 9b; blue) cortex. There was a very strong contra lateral bias for right spatial positions as the recording was done in the left hemisphere. There were relatively large numbers of position neurons in parietal cortex as compared to prefrontal cortex in general. Since most neurons preferred right spatial positions we had more neurons contributing to the left half of the population tuning curve aligned to the peak response in each cell. Consequently, the left part of the curve was smoother (Fig. 9a).

Comparison of neural signals coding boundary position in parietal and prefrontal cortex

Similar to position neurons above, we found boundary neurons that were activated by the presentation of the boundary cue at one particular location in the display both in parietal and prefrontal cortex. The DYSC-shift task was designed so that the spatial category of the sample was defined based on two different spatial variables - the spatial location of the sample stimuli and the spatial location of the boundary. Presenting the two spatial stimuli at different times in the trial helped us to dissociate the neural signals that were encoding them. Rasters and spike density functions in Figure 10 illustrate a parietal neuron (a-d) and a prefrontal neuron (e-h) that were activated preferentially by the presentation of the boundary at one or the other location in the display (either 5° to the left or right of the gaze fixation target). The parietal neuron exhibited higher activity

when the boundary was presented at the left boundary location (Fig. 10a, b) than when it was presented at the right boundary location (Fig 10 c, d), whereas little difference in activity was evident as a function of whether the sample was assigned to the LEFT (Fig. 10a, c) or RIGHT spatial category (Fig. 10b, d).

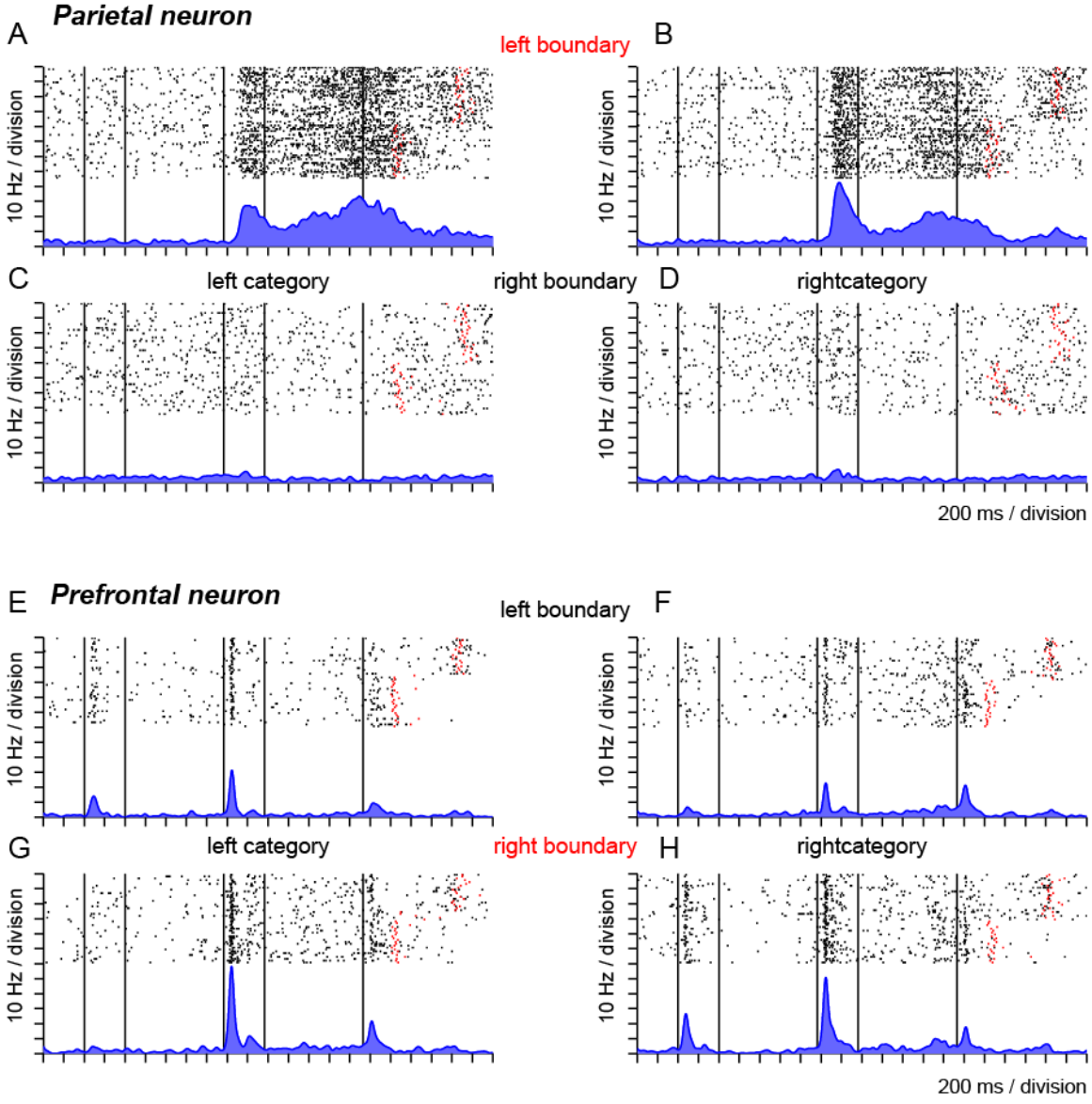

Figure 10. Rasters and spike density functions of a parietal (***A-D***) and a prefrontal neuron (***E-H***) encoding boundary location. Trials are divided into four types as a function of spatial category and boundary position. (*A, B, E, F*) Left boundary trials.

(*C, D, G, H*) Right boundary trials. (*A, C, E, G*) Left category trials. (*B, D, F, H*) Right category trials.

Similarly, the illustrated prefrontal neuron exhibited greater activity when the boundary cue was presented at the right spatial location (Fig 10 g, h) compared to when boundary was presented at the left spatial location (Fig. 10 e, f), with little obvious variation in firing rate as a function of category (compare Fig. 10e,g to 10f, h). These results indicate that there are neurons in both parietal and prefrontal cortex which specifically code the spatial location of the boundary cue applied to divide the spatial array into categories.

To visualize the time course of population activity coding boundary position, we constructed average normalized spike density functions separately using subsets of trials when the boundary was presented at the preferred position (Fig. 11; solid SDF) and the non preferred position (Fig. 11; dashed SDF) for each neuron for neural populations in parietal (Fig 11a) and prefrontal cortex (Fig 11b). The parietal population exhibited more pronounced neural activity coding boundary position with a peak during the boundary period, while the prefrontal population activity was weaker.

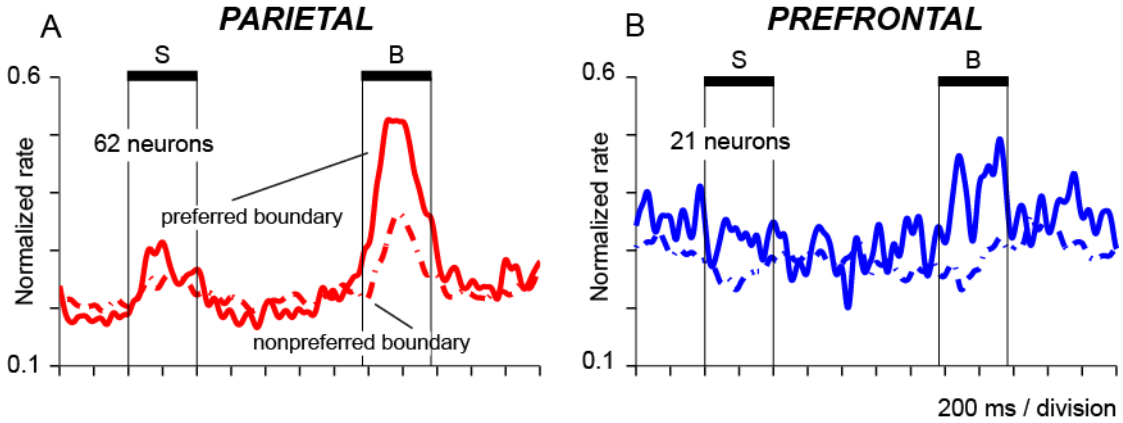

Figure 11. Population average normalized spike density functions illustrate the time course of activity for boundary neurons in (***A***) parietal cortex and (***B***) prefrontal cortex. Trials are divided into preferred (solid SDF) and non preferred boundary positions (dashed SDF).

Comparison of neural signals coding spatial category in parietal and prefrontal cortex

In the above sections, we described the presence of neurons that code the spatial position of the sample and the boundary. To evaluate the influence of spatial category on the firing rate of the neurons, we performed a 2-way ANCOVA on firing rate in the boundary and delay-2 periods; with spatial category and boundary location as the two factors (we included the rate of firing during the prior sample and delay periods as a covariate in the analysis). We defined pure category neurons as those in which firing rate varied only as a function of the spatial category factor and not boundary location. Conversely, neurons in which firing rate varied as function of the interaction between category and boundary factors exhibited significantly higher firing rate for one category (left or right) than the other, but the category signal was significantly modulated by boundary location. We found 185 neurons total in parietal and prefrontal cortex in both monkeys that encoded the spatial category of the sample ($p < 0.05$ for the category factor

and not significant for boundary or interaction). We provide examples in Fig 12 in which we illustrate the activity of two different neurons (one in parietal and other in prefrontal cortex) that encode spatial category. On right boundary trials, the parietal neuron shown in Fig 12a-d, was significantly more active during the boundary and delay-2 periods whenever the sample was assigned to the left spatial category (Fig. 12c) in comparison to the right spatial category (Fig. 12d), whereas little or no category-selective activity was evident on left boundary trials (Fig. 12a, b) Note that this should not be confused as a signal reflecting the boundary location alone as the firing rate of the neuron differs as a function of the spatial category across trials with the same boundary position (Fig 12c and d). Moreover, this signal cannot simply reflect a retinocentric preference for a particular subset of sample positions, because samples presented at the same intermediate locations in the grid, between the boundary locations, could either activate this neuron (Fig. 12c) or not (Fig. 12b) depending on the relationship of those positions to the boundary. The illustrated prefrontal neuron had higher activity when the sample belonged to the right category (Fig. 12f, h) than when it belonged to left category (Fig 12e, g). The category signal in this neuron was mostly confined to the end of the boundary presentation and beginning of the subsequent delay 2 periods. Both neurons were significantly influenced by the category factor and the interaction between the category and boundary factors, indicating that the strength of the category signal varied significantly between the two boundary positions. These signals, although weak, were also characteristic of the population, in which we saw relatively weak category signals

distributed among a relatively large number of neurons, so that category-selectivity was perhaps more strongly evident at the population level (Fig. 13)

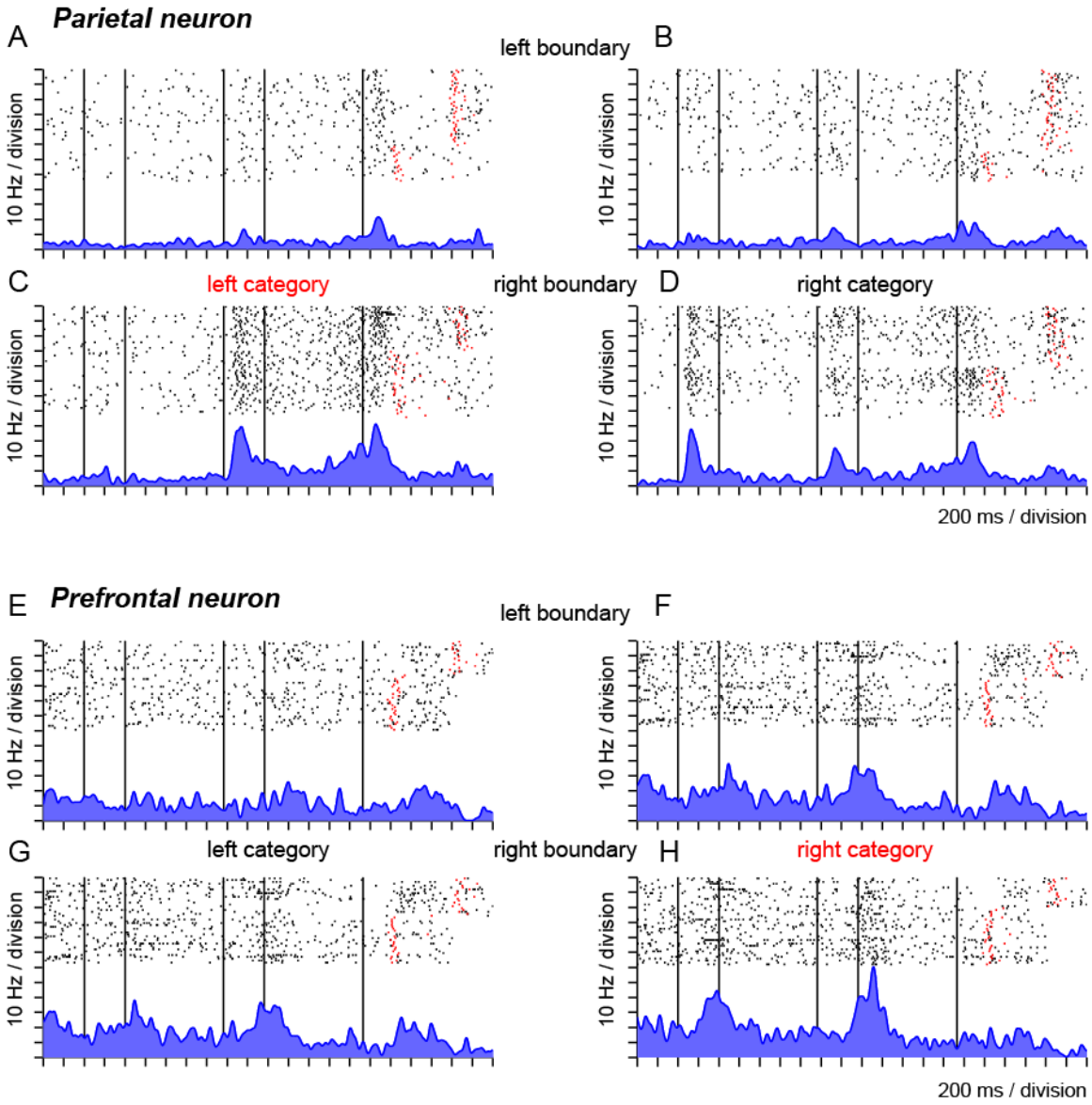

Figure 12. Rasters of a parietal (***A-D***) and a prefrontal neuron (***E-H***) representing spatial category information. Trials are divided into four groups as a function of spatial category and boundary position. (***A, C, E, G***) LEFT category trials. (***B, D, F, H***) RIGHT category trials. (***A, B, E, F***) Left boundary trials. (***C, D, G, H***) Right boundary trials.

To evaluate the time course of category signals in parietal (Fig. 13a) and prefrontal (Fig. 13b) cortex at the population level, we plotted mean population spike density functions (sdf) of category-selective activity in the two areas. For this analysis – we focused on pure category neurons, including neurons in the population that only showed a significant relation to the category factor and had no significant effects for boundary position or the interaction of category with boundary position. The signal coding the category of the sample dissociated from boundary position was found to be distributed between parietal and prefrontal cortex.

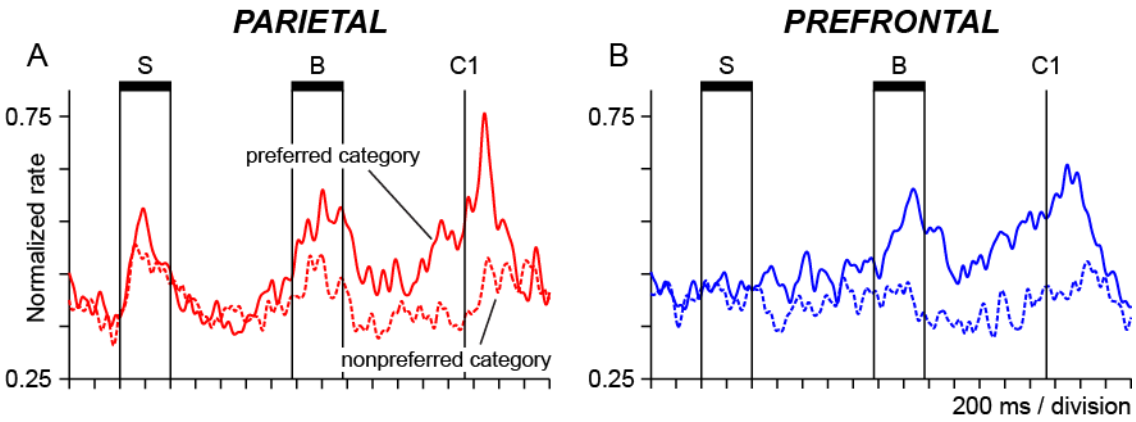

Figure 13 Population spike density functions illustrating the time course of category-selective activity in parietal cortex (***A***) and prefrontal cortex (***B***). Activity on preferred category trials is plotted with a continuous line whereas activity on non-preferred category trials is plotted with a dotted line. The vertical lines in the plot represent the onset/offset of various stimuli. From left to right, the first two vertical lines indicate sample onset and offset (horizontal black bar marked 'S'). The next two lines indicate onset/offset of the boundary (black bar marked 'B'). The last vertical line indicates choice-1 onset.

Decoding sample position, boundary position, and spatial category from population activity in parietal and prefrontal cortex

We applied pattern classification analysis to population activity in parietal and prefrontal cortex in order to decode the basic spatial variables of the task: horizontal sample position, horizontal boundary position, and spatial category (LEFT or RIGHT).

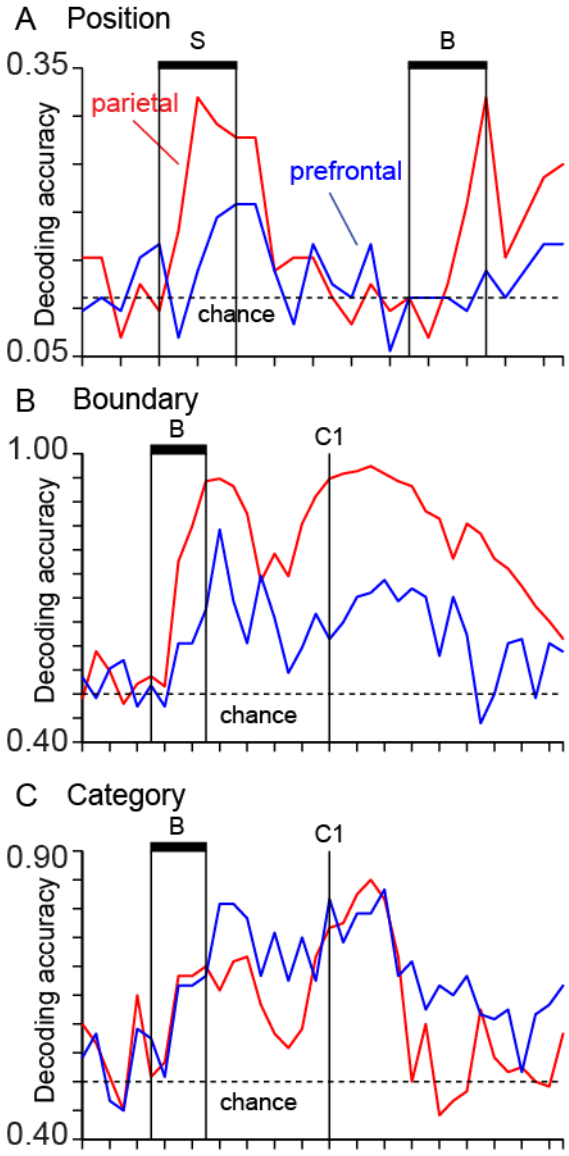

Figure 14. Time courses illustrate fluctuations in decoding accuracy (proportion of trials correct) when decoding (***A***) the spatial position of sample, (***B***) boundary position, and (***C***)

the spatial category of the sample based on population activity patterns in parietal cortex (red) and prefrontal cortex (blue).

We performed this analysis using successive 100 ms time bins of population activity in parietal and prefrontal cortex. The proportion of trials at each time point that the decoding analysis returned the correct value for sample position, boundary position, or spatial category (given the stimuli presented each trial) provided a way to measure the time course with which population activity in parietal and prefrontal cortex represented each task variable. When decoding the spatial position of the sample stimulus, decoding accuracy increased first and reached a greater peak level in parietal cortex (Fig. 14a; red) relative to prefrontal cortex (Fig. 14a; blue; decoding accuracy predicted by chance is 11% correct given that sample stimuli were displayed at 9 different horizontal positions), consistent with the hypothesis that visual information flows forward from parietal to prefrontal cortex. The decoding accuracy time course for the location of the boundary followed a similar pattern as that above, increasing earlier and reaching greater levels of accuracy in parietal cortex relative to prefrontal cortex (Fig. 14b). The most interesting result related to the population coding of spatial category (which is an abstract cognitive property), for which we found the converse pattern. The decoding accuracy time course for category reached higher levels when based on population activity in prefrontal cortex as compared to parietal cortex (Fig. 14c), although the difference was modest. In both cortical areas, the population signal coding category (Fig. 14c) increased after the appearance of the boundary, at the time that the two stimuli (sample and boundary) that jointly define spatial category had been presented.

Position invariance of category signal

To examine in greater detail the characteristics of the category signal, we evaluated its variance with respect to stimulus position, quantifying the degree to which the category signal varied in strength as a function of whether the boundary was presented at the left or right locations in the display (shifting the position of categorized stimuli accordingly). This determined the degree to which the category signal in single neurons was 'pure' in the sense of being truly invariant with respect to retinal position, or varied in strength as a function of retinal position. Our prediction was that there would be some 'position invariant' category cells in which the category signal did not depend on the position of the boundary or sample, while there would be other cells which in which the category signal was modulated by the position of sample or of the boundary.

To address this question, we examined the spatial tuning of category neurons separately for left and right boundary trials. If the category signal was position invariant, the spatial tuning functions of category neurons would shift with the boundary position and the modulation of activity between preferred and nonpreferred categories would be about the same in either case.

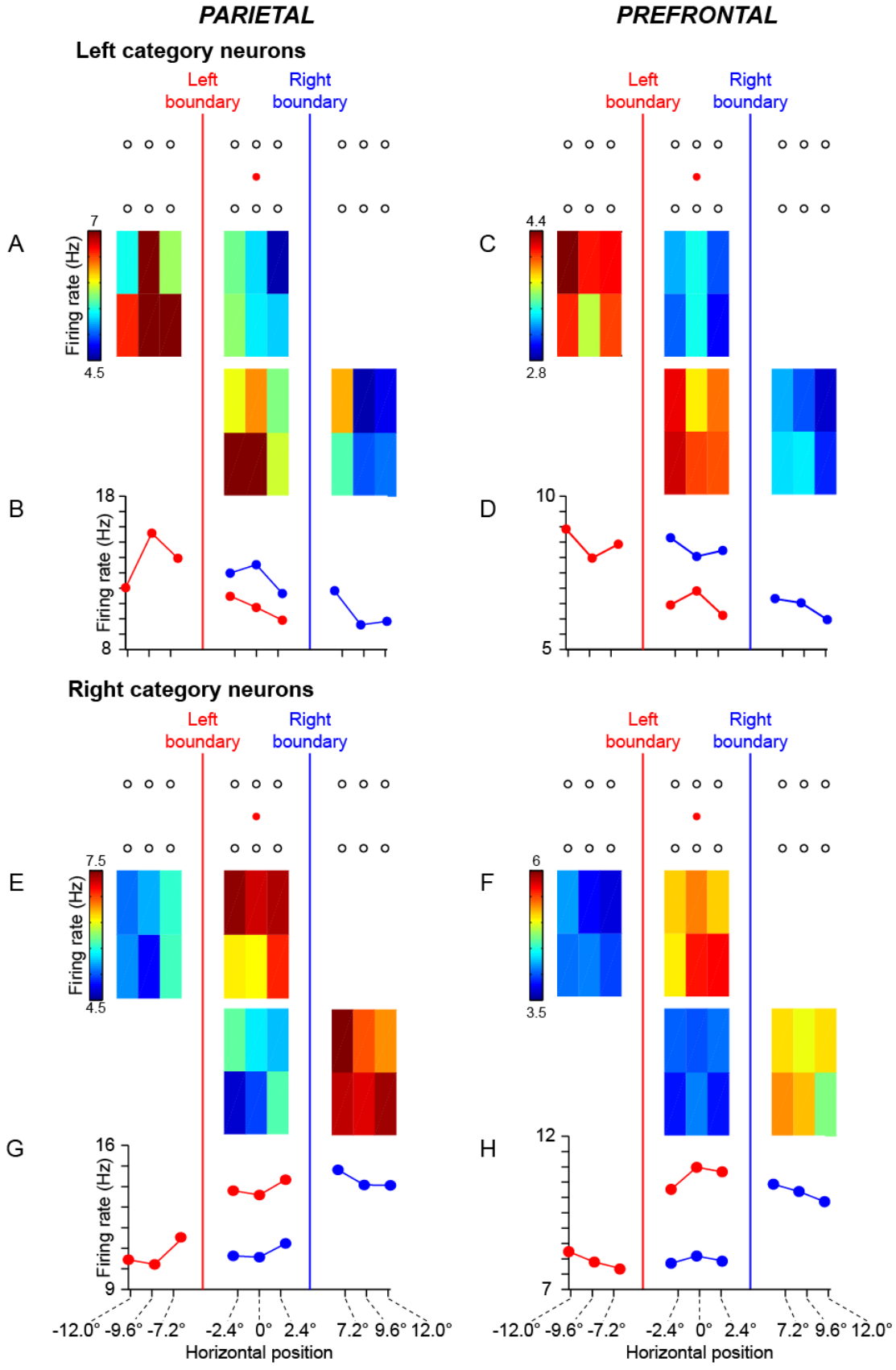
PARIETAL
PREFRONTAL
Left category neurons
Left boundary
Right boundary
A
B
C
D
Firing rate (Hz)
7
4.5
18
8
4.4
2.8
10
5
Right category neurons
E
F
G
H
7.5
4.5
6
3.5
16
9
12
7
-12.0° -9.6° -7.2° -2.4° 0° 2.4° 7.2° 9.6° 12.0°
Horizontal position

Figure 15. Spatial tuning plots for LEFT-preferring (*A-D*) and RIGHT-preferring (*E-H*) category neurons in parietal (*A,B,E,G*) and prefrontal cortex (*C,D,F,H*). ***A,C,E,F –*** Color grids illustrate the mean population firing rate at each sample position separately for the two boundary locations (warmer colors indicate higher firing rate, see scale at upper left of each panel). ***B,D,G,H***– Population spatial tuning curves illustrate firing rate at different horizontal sample positions, plotted separately for the two boundary locations (left: red and blue: right).

Importantly, if the category tuning functions are truly position invariant then the tuning plots for the sample positions that lie in between the two boundaries will reveal that different levels of neural activity are associated with each sample position depending on the relative location of the boundary (for samples located in between the boundary positions, the spatial category of the sample will flip from LEFT to RIGHT or vice versa depending on the relative position of the boundary). Figure 15 illustrates the population spatial tuning of category-selective neurons (Pure category neurons restricted to neurons with NO significant boundary or boundary X rule interaction ANOVA, $p < 0.05$; Fig. 15a-d, LEFT preferring category neurons; Fig. 15e-h, RIGHT preferring category neurons) plotted separately for left and right boundary locations. Each box in the color grids represents the mean population firing rate for a different stimulus location in the array. (There were 12 different sample positions per boundary, and the sample arrays partially overlapped at the two boundary locations giving 18 unique sample positions total; see diagrams of the spatial locations of the fixation target, boundary locations, and sample positions at the top of Fig. 15a, c, e, f.) Color represents the intensity of firing (red indicates higher rate and blue indicates lower rate of firing). The spatial tuning curves underneath the color grids indicate the mean firing rate at each sample x position (collapsing across the two sample y positions; Fig 15b, d, g, h). There is some variation

in the population activity across various sample positions within a category (more so for left category neurons; Fig 15a-d), however the spatial category preference of neurons was maintained across different boundary positions and absolute sample positions in the display. For example right category neurons (fig 15e-h) exhibit higher firing rates when the sample was located to the right of the boundary whether the boundary appeared at the left position (red) or the right position (blue) in the display. This indicates that at the population level, neurons maintain their spatial tuning characteristics for categories across different spatial positions.

Because the sample positions in between the two boundaries were alternatively assigned to the LEFT and RIGHT spatial categories we were able to confirm that category-selective activity did not depend on the retinal position of the sample, in the sense that we found that the samples presented at the same position in space elicited different levels of neural activity in the population based on the spatial relationship of that position to the boundary (note that portions of red and blue spatial tuning curves in between the two boundary positions don't overlap indicating that these same sample positions were associated with different levels of population activity in category neurons; Fig 15b, d, g, h). This point is of significant importance as it indicates that neurons modified their spatial tuning for sample position based on the relationship to the boundary, providing evidence that at the population level, the spatial tuning of these cells was dependent on the spatial category of the sample irrespective of the spatial position.

Note that the above pattern of position invariance in the spatial tuning of category neurons at the population level could be produced by summing the activity of single neurons that maintained their category preference between left and right boundary locations but exhibited stronger category signals for one or the other. Therefore, to quantify the influence of boundary location, spatial category and in particular their interaction on the activity of single neurons we measured the firing rate of each neuron (without pre-selection) within a sliding 200 ms window advanced in 20 ms steps, and regressed firing rate over trials at each time step onto category, and the interaction between category and rule. In this analysis, 'pure' category neurons would be those with activity that was strongly related to spatial category, but not related to boundary position, or the boundary-category interaction. In contrast, neurons with activity significantly influenced by the interaction between category and boundary would exhibit a category signal that varied in strength depending on whether the boundary was left or right, and this sense would reflect a type of category signal less completely dissociated from position information. The regression analysis produced a time series of regression coefficients for each neuron showing the strength of the relation between firing rate and boundary location, spatial category, and the interaction between category and boundary, at each time step (Fig. 16). We found by this analysis that population activity tended to exhibit a stronger relation to spatial category in prefrontal cortex (Fig. 16h), whereas population activity tended to exhibit a stronger relation to boundary position (Fig. 16g) and the interaction between category and boundary position (Fig. 16i) in parietal cortex, based on the mean population regression coefficient averaged across all neurons in the

population at each time step. Ranking neurons according to the time of the peak coefficient, early to late from bottom to top in Figure 16a-f (each row represents data from one neuron), produced a diagonal band of color in each plot indicating the rate at which neurons were recruited to represent boundary, category or their interaction in each cortical area.

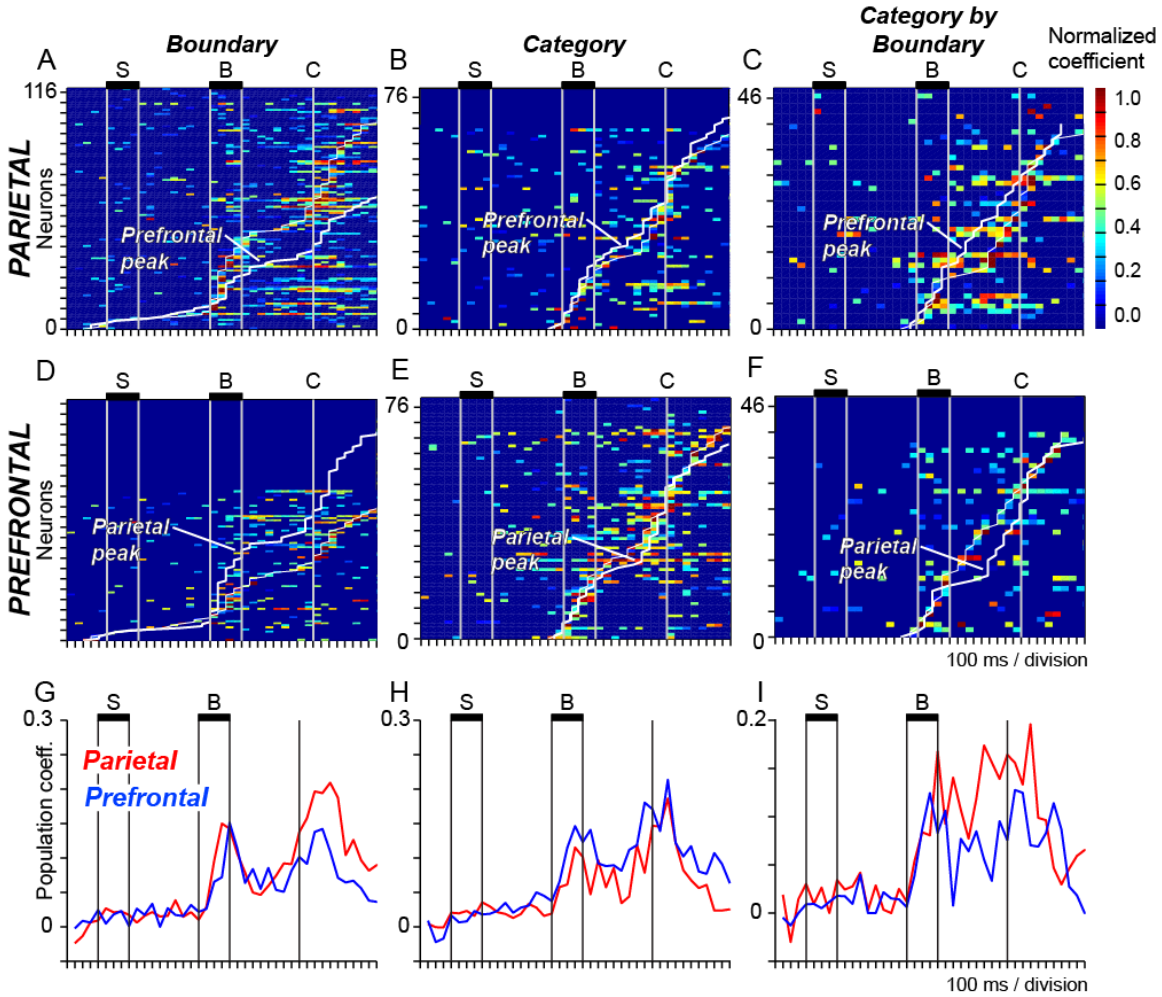

Figure 16. Time course of regression coefficients obtained for boundary, category and the interaction between category and boundary for individual neurons in parietal and prefrontal cortex. Each row in the color plots indicates the time series of regression coefficients obtained for a single neuron normalized to the peak coefficient for that neuron. ***A, D.*** Regression coefficients for boundary obtained for neurons in parietal cortex *(A)* and prefrontal cortex (*D*) ***B,E.*** Regression coefficients for spatial category obtained for neurons in parietal cortex (*B*) and prefrontal cortex (*E*) ***C,F.*** Regression

coefficients for the interaction between category and boundary obtained for neurons in parietal cortex (*C*) and prefrontal cortex (F). White lines show the time of peak coefficients in the other cortical area for comparison. Dark blue regions indicate time points at which coefficients were non-significant. ***G-H***. Time courses plot the mean, non-normalized population regression coefficient obtained at each time point in parietal (red) and prefrontal (blue) cortex for boundary position (*G*), category (*H*) and the interaction between boundary position and category (*I*).

Pairs of plots illustrating recruitment in parietal and prefrontal cortex for a given variable (for example Figs. 16a and d) have the same scaling along the y-axis (indicating numbers of neurons) in the two areas so the slopes of recruitment can be directly compared. Neurons with activity reflecting spatial categories were recruited at a faster rate or at earlier times in the trial in prefrontal cortex in comparison to parietal cortex (Fig. 16e, the diagonal band of color indicating prefrontal recruitment leads, albeit slightly, the white line indicating parietal recruitment), whereas the neurons with activity reflecting boundary position were recruited faster and at earlier times in parietal cortex (Fig. 16a, d). Neurons reflecting the interaction between category and boundary were first engaged in prefrontal cortex (Fig. 16c and f) although as noted above, the strength of the effect as measured by the mean population regression coefficient, was stronger in parietal cortex (Fig. 16i).

In continuation of the above set of findings suggesting that the effect of the interaction between boundary and category was stronger in parietal cortex (indicating weaker position invariance and therefore a less 'pure' category signal), we evaluated the results of the prior 2-way ANOVA (with sample side and boundary location as factors) in terms of the number of 'pure' category neurons (significant for category but not the

interaction with boundary) and the number of position-dependent category neurons (significant for both category and the interaction with boundary) in the two areas – parietal and prefrontal cortex (Fig 17). Blue bars represent the number of neurons that were significant for the category-boundary interaction (position-dependent category neurons) while red bars represent the number of neurons that were significant for category and not the interaction with boundary ('pure' category neurons).

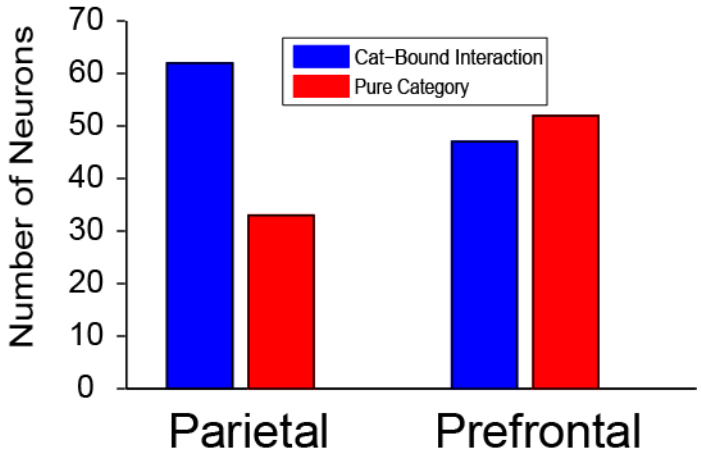

Figure 17. Histograms indicate the number of neurons in parietal and prefrontal cortex with activity significantly related to both category and the category-boundary interaction (blue bars), as well pure category neurons (red bars) with activity significantly related to only the category factor in the analysis.

The results of this analysis indicated that a larger fraction of category neurons exhibited position invariance in prefrontal relative to parietal cortex (Fig. 17). Specifically, we found that whereas 53% of all category neurons in prefrontal cortex were position invariant by the above definition, only 34% of all category neurons in parietal cortex were position invariant. Category is an abstract property (that is, no specific stimulus represents left or right, rather category is defined by the spatial relationship between two different stimuli). The observation that a greater number of neurons in

prefrontal cortex represented pure category relative to parietal cortex is consistent with our hypothesis that prefrontal cortex is more predominantly involved in higher level spatial cognition. These numbers between parietal and prefrontal were significantly different (t-test; p=0.014).

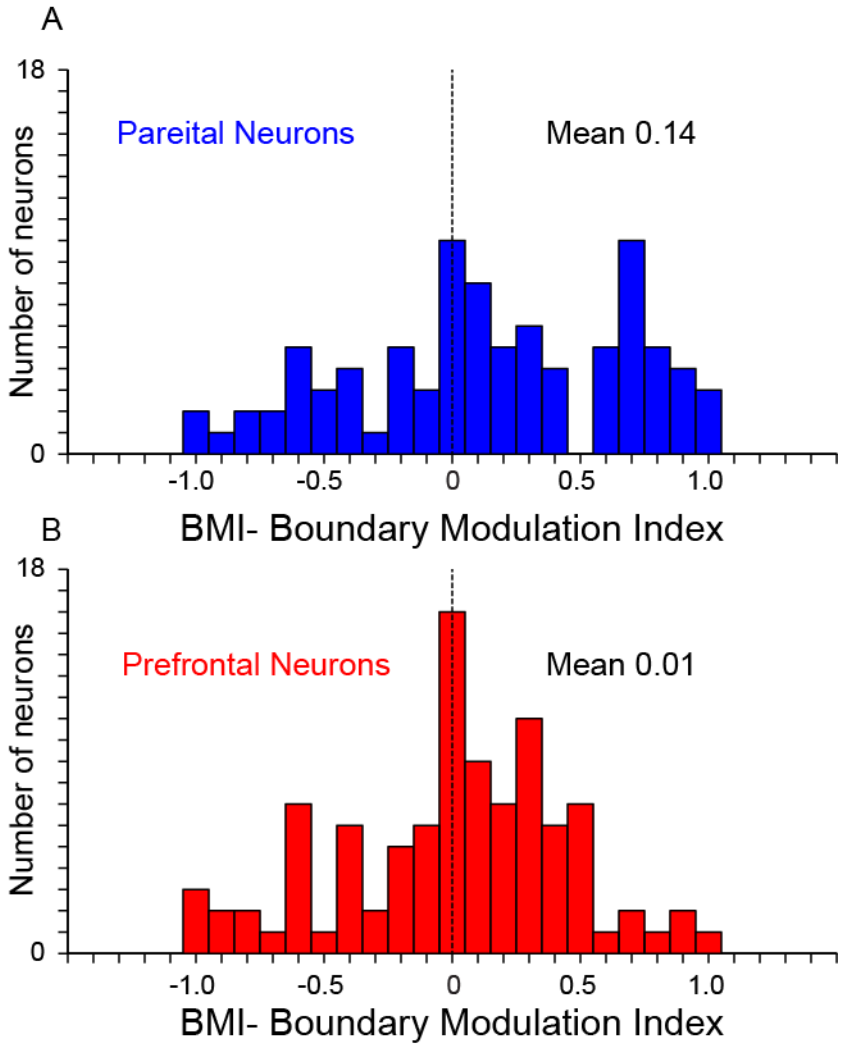

Figure 18. Frequency distribution histograms of the boundary modulation index (BMI) for (A) parietal neurons and (B) prefrontal neurons.

To further quantify the influence of boundary position on the category signal, we calculated an index we termed the boundary modulation index (BMI). We calculated the

BMI index for all category selective cells ($p < 0.05$ for category factor in the ANOVA) in both parietal (Fig 18a) and prefrontal (Fig 18b) cortex. To calculate the index, we utilized firing rates during the stim 2 and delay 2 periods on trials when the sample was assigned to the preferred category for each neuron. We defined the congruent boundary as the one on the same side as the category preference for each neuron (for example, the left boundary was congruent for LEFT-preferring category neurons). We then computed the difference in firing rate on preferred and nonpreferred category trials for the two boundary location, and then defined the BMI as the difference between these numbers divided by their sum. The value of this index could vary from -1 to +1 (where -1 indicates greater activity on incongruent boundary trials and +1 the reverse). A BMI value of 0 indicates no difference in preferred category activity between the two boundary positions tested, and therefore position invariance in the category signal. First we note that the mean BMI for both cortical areas was centered near 0, however the values were broadly distributed between +1 and -1 indicating a considerable degree of boundary dependence in some neurons in both areas. Second, we found that the BMI index had a broader distribution with a larger positive tail in parietal cortex relative to prefrontal cortex, with a correspondingly larger mean value in parietal cortex (0.14) relative to prefrontal cortex (0.01). This indicates a stronger tendency for BMI indices to have absolute values closer to 0 in prefrontal cortex than in parietal cortex indicating stronger position invariance. Thus the BMI index gives us some insight into the differences in spatial category coding in parietal and prefrontal cortex.

Physiological interaction between position and category neurons

We next wanted to look at the potential physiological interaction between category and position neurons. As a first approach to this question, we evaluated the degree to which baseline activity was correlated in pairs of neurons in which one neuron coded sample position and the other neuron coded spatial category, as this could potentially reflect a synaptic interaction between the neurons (most likely through indirect pathways). In Fig. 19, we illustrate the number of position and category neuron pairs with significantly correlated baseline activity ($p < 0.05$) as well as the mean cross correlation coefficient obtained in parietal and prefrontal cortex. We performed separate analyses in which the position and category neuron had the same spatial preference (Fig. 19; blue bars) versus different spatial preferences (Fig. 19; red bars). We restricted this analysis to position neurons that preferred positions in the outermost thirds of the sample array (as these sample positions were associated with only one spatial category) and then defined a neuron pair as having the same preference in the case that, if the position neuron preferred positions in the leftmost third of the sample grid, for example, the category neuron preferred the LEFT spatial category, and vice versa. We also repeated the analysis using pairs of neurons that were simultaneously recorded (in which case correlated baseline activity could reflect real-time physiological interaction), and pairs of neurons that were recorded at different times (providing a measure of the degree of spurious correlation due to other factors).

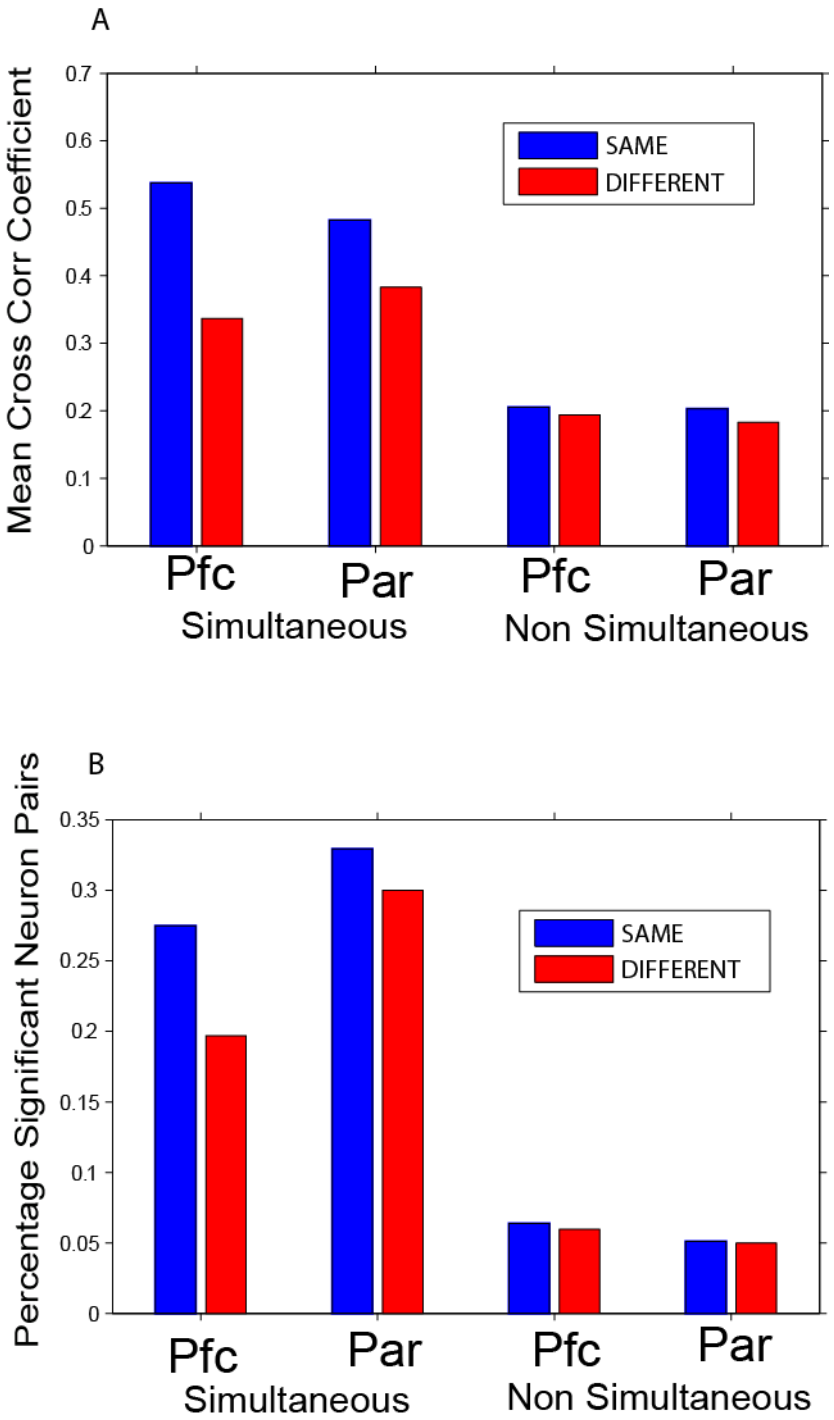

Figure 19. Analysis of correlated baseline activity (in the fixation period before sample onset) over trials in pairs of neurons in which one neuron coded sample position and the other sample category. ***A***. Mean cross correlation coefficient as a function of whether the position and category neuron in each pair had the same spatial preference (blue bars) or opposite spatial preference (red bars) in prefrontal cortex ('Pfc') and parietal cortex ('Par'). Data are plotted separately for neuron pairs recorded simultaneously (left) and at

different times (right). ***B***. The percentage of neuron pairs with significantly correlated baseline activity.

First, we found significantly correlated baseline activity in approximately 20-30% of neuron pairs in the case that the two cells were recorded simultaneously, versus approximately 5% of neuron pairs recorded at different times (Fig. 19; the expected false-positive rate given the 0.05 alpha level of the test). Second, we found that correlation coefficients were slightly stronger on average (Fig. 19, top), and more neuron pairs exhibited significantly correlated activity (Fig. 19, bottom) in the case that position and category neuron in each pair had the same spatial preference (Fig. 19; blue bars) versus different spatial preferences (Fig. 19; red bars).

Discussion

The Dynamic Spatial Categorization (DYSC) was designed to dissociate feature from abstract property (i.e. category on our case) signals at a single cell level in both prefrontal and parietal cortex. In order to do so, we shifted the location of the category boundary requiring monkeys to assign the same sample position to opposite spatial categories over trials, which made it possible in turn to separate information about the absolute position of the sample from its spatial category. In addition, in the DYSC-shift experiment we presented first the sample stimulus followed after a delay by the boundary cue. Thus, a key feature of the DYSC-shift task was that the monkey had to combine information stored in spatial working memory (the sample position) with information provided by visual input (the boundary cue) in order to compute the spatial category of

the sample. This further helped to dissociate neural signals coding the spatial category of the sample (evoked presumably by the appearance of the boundary cue) from neural signals coding the position of the sample (evoked earlier in the trial by the presentation of the sample). Based on the SHIFT aspect of the task, the sample positions that were in the middle location between the two boundary locations could be classified as left or right depending on the boundary positions. This aspect of the task design helps us to decipher category based purely on spatial location from that of the category based on the spatial relationship between two objects.

Prior studies have shown that very similar distributed patterns of activity are found in parietal and prefrontal cortex during tasks that test spatial working memory, such as the ODR (Oculomotor Delay Response) task (Chafee and Goldman-Rakic 1998). One of the motivations behind our study was to look into the distribution of neuronal activity patterns between the two areas in the case that the complexity of the task was increased. We also found that there were similar patterns of activation related to the DYSC-shift task in the two areas – neural signals reflecting sample position, boundary position, and spatial category were distributed both to parietal and prefrontal cortex (Figs. 3, 6-18). In both cortical areas, we found that population spatial tuning functions of category-selective neurons (plotted with respect to the retinal position of the sample) shifted with the position of the category boundary, showing that activity was not an obligatory function of absolute sample position or absolute boundary position (Fig. 15). These data suggest that parietal, and also prefrontal, neurons code relative spatial

position, consistent with the prior finding of object-centered coding in parietal cortex (Crowe, Chafee et al. 2004; Crowe, Averbeck et al. 2008). These data are the first demonstration that both parietal and also prefrontal neurons can represent spatial categories based on spatial relationships.

However we were also able to detect differences between parietal and prefrontal cortex. We found that the neural signal coding the spatial position of the sample and that of the rule appeared earlier and was stronger in parietal cortex in comparison to prefrontal cortex. On the other hand, we found prefrontal cortex contained a larger number of category-selective neurons that parietal cortex (Fig. 3), and that the spatial category signal was stronger in prefrontal cortex in comparison to parietal cortex (Figs. 14, 16). This order of activation in the network, with prefrontal leading parietal in the representation of spatial categories, will be more fully addressed in Chapter 2, and is similar to the results obtained by Merchant and colleagues (Merchant, Crowe et al. 2011) but is opposite to the order of spatial category representation reported by Swaminathan and colleagues (Swaminathan and Freedman 2011) who found in a spatial categorization task that parietal cortex lead prefrontal cortex in category representation. It is hard to identify the actual cause of this discrepancy, but one factor that differed between the three studies was the area of recording in the parietal cortex. Our study and the Merchant et al. study (Merchant, Crowe et al. 2011) recorded activity in Area 7a, whereas Swaminathan and colleagues recorded neural activity in area LIP (Swaminathan and Freedman 2011), and it is quite possible that these two areas play different roles in categorization. Another potential difference is that our study required that monkeys use

working memory to categorize the stimuli, and spatial categories were dissociated from spatial features in the task, which was not the case in the Swaminathan study, in which the direction of visual motion and the spatial category were correlated.

In addition, we sought to determine the degree to which category signals in parietal and prefrontal neurons were pure in the sense of being dissociated from spatial position, and found that category-selective neurons exhibited a varying degree of translation invariance with respect to boundary position. We found that at least in some neurons, activity could code spatial category independently from spatial position (even though in many neurons these signals were relatively weak), and that there were more of these neurons in prefrontal cortex, where neurons showed a greater degree of translation invariance relative to parietal neurons (Figs. 17, 18), suggesting prefrontal cortex plays a more direct role in abstract category representation.

We also found that different populations of neurons were active to encode the spatial position of the sample at different times in the trial: the sample period, the delay period, and also the boundary period, in both parietal and prefrontal cortex (Fig. 7). This could reflect the need to carry information about sample position forward in time to provide a basis for comparison between sample and boundary that is required to compute spatial category. Another interesting feature of our results is that we found that the baseline activity of position-selective neurons and category-selective neurons was correlated over trials, but more strongly so if the preferred position of one neuron fell within the preferred category of the other and only if the neurons were recorded at the same time (Fig. 19). This revealed a differential interaction between position and

category neurons based on whether their spatial preferences were congruent or incongruent; a finding that suggests that position signals may drive category signals. Past studies from our lab have showed a similar relationship between correlated baseline activity and spatial coding in parietal cortex, with neurons coding the same object-centered position having more strongly correlated activity than neurons that code different positions (Crowe, Averbeck et al. 2010).

In DYSC-shift monkeys assigned positions to categories according to one rule (whether the horizontal coordinate of the sample fell to the left or right of the boundary). To look into neural mechanisms that the brain uses to categorize stimuli according to different rules, and therefore to investigate then neural mechanisms of executive control over categorization, we designed the DYSC–rotate task. This task will help us to isolate neural signals in the network that were selective for spatial categories and were also rule-dependent, to provide a cellular correlate of executive control over a cognitive process.

Chapter -2

Executive control of spatial cognition in prefrontal-parietal network

Introduction

Executive control refers to the brain's ability to flexibly process input information conditioned by different internal state variables, such as rules or goals, to produce different actions as a function of behavioral context. Prefrontal cortex is involved in executive control and this it has been shown in humans as well as monkeys (Goldman-Rakic 1987; Wallis, Anderson et al. 2001; Miller, Freedman et al. 2002; Nakahara, Hayashi et al. 2002; Genovesio, Brasted et al. 2005; Mansouri, Matsumoto et al. 2006; Stoet and Snyder 2009). Although much research has been done in executive control processes, the cellular and neural mechanism still is largely unknown.

The DYSC-Rotate task provides an experimental approach to investigate the neural mechanisms of executive control specifically over cognitive processing in distinction to sensorimotor processing. This task uses a variable rule to control how stimuli are assigned to spatial categories based on different spatial relationships applied to differentially group the stimuli. The spatial nature of the stimuli as well as the categories defined by the task implies that the dorsal stream will be importantly involved. Therefore we chose to record from prefrontal and posterior parietal cortex simultaneously to determine how the processing demands of the task might be shared between or differentially subserved by the two cortical areas, allowing us to specifically test the hypothesis that PFC is more fundamentally important to the neural and network basis of

executive control. Based on previous research from our lab (Chafee, Crowe et al. 2005; Chafee, Averbeck et al. 2007; Crowe, Averbeck et al. 2008), it has been shown that parietal cortex is involved in object-centered spatial coding, evidence that this cortical area participates in the representation of spatial relationships which provides the basis for spatial categories in the present experiment. Also, the work of Apostolos (Merchant, Crowe et al. 2011) and Miller (Jo, Freedman et al. 2012; Palmer, McDonough et al. 2012) has shown involvement of prefrontal cortex in spatial categorization. However, note that none of the prior studies have investigated the neural mechanisms of executive control over categorization.

Neural representations of rules have been shown to be present in both parietal (Stoet and Snyder 2004; Stoet and Snyder 2009) and prefrontal cortex (Wallis, Anderson et al. 2001; Muhammad, Wallis et al. 2006), but the focus of the present experiment is how rules modulate other signals related to cognitive processing, and this has not been previously studied at the single neuron level in this network. Thus, to characterize the neural mechanisms involved in computational flexibility and executive control over spatial cognitive processing at the network level, in our task, monkeys applied different criteria to flexibly regroup the same set of stimulus positions into alternative spatial categories, placing categorization as a cognitive process under executive control. Our hypotheses for the DYSC-Rotate task are 1) Neural signals coding rule-dependent category will be distributed in the parietal-prefrontal network. 2) Rule-dependent category signals will be stronger and earlier in prefrontal cortex than parietal cortex

(potential generator of executive control at network level) and 3) Rule-dependent category signals will be transmitted from prefrontal to parietal cortex.

METHODS

Animals Two male rhesus macaque monkeys (5-8kg) were trained to perform the DYSC-rotate task (below). Neural recording was performed simultaneously from prefrontal area 46 and posterior parietal area 7a. All surgical and experimental procedures associated with surgery and neurophysiological recording conformed to NIH Guidelines and protocols approved by the Internal Animal Care and Use Committee (IACUC) at the VA Medical Center and the University of Minnesota. We prepared monkeys for neural recording in an aseptic surgery performed using gas anesthesia (Isoflurane, 1-2%). We placed a craniotomy centered over the principal sulcus (Brodmann's area 46) in the left prefrontal cortex, and another craniotomy centered over the posterior aspect of the inferior parietal gyrus (Brodmann's area 7a) in the left posterior parietal cortex, of both monkeys (Fig. 20d, e). We placed titanium screws in the skull, and attached recording chambers (either 7 or 13 mm i.d.) and posts for head fixation to the screws using surgical bone cement. We administered post-surgical analgesia for several days (Buprenex, 0.05 mg / kg Bid, i.m.).

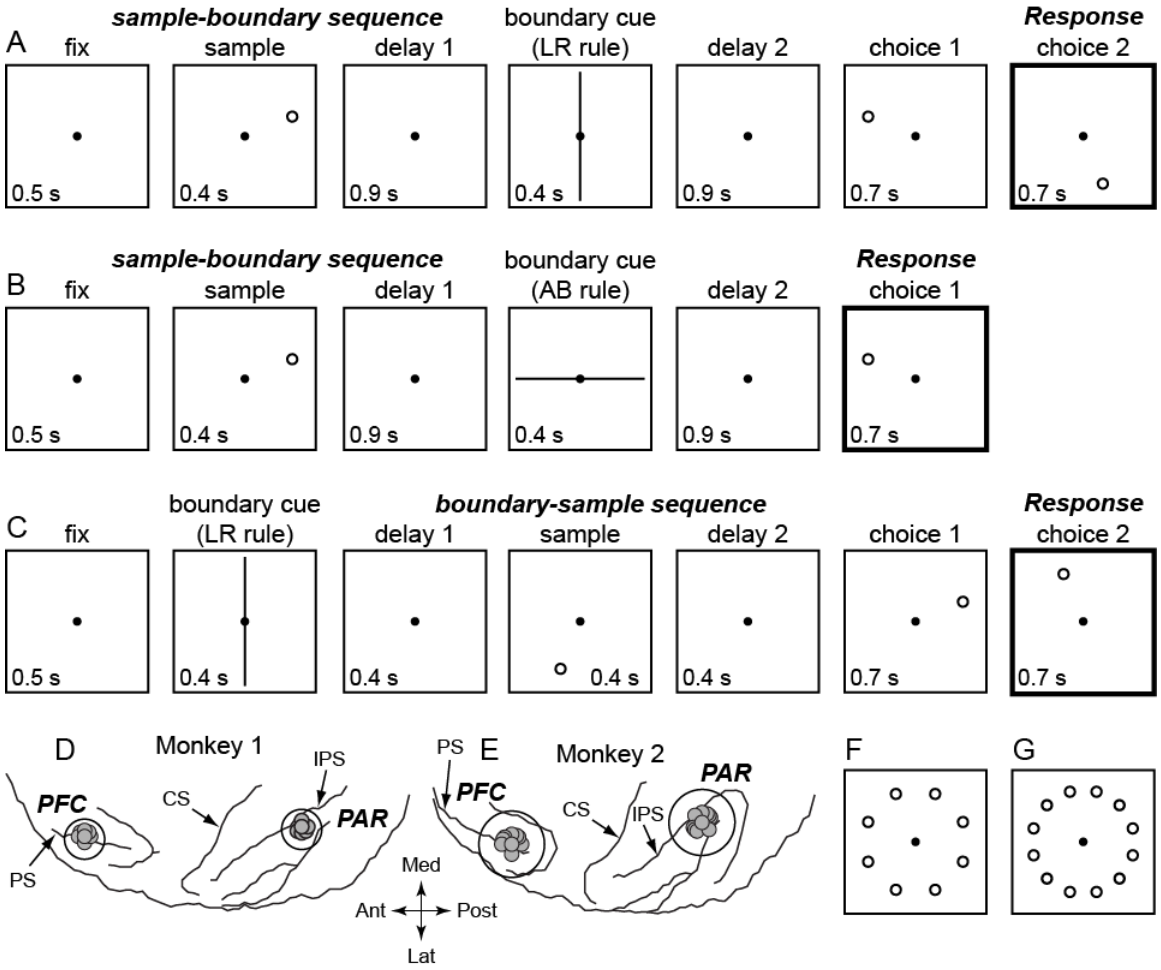

Figure 20. Event sequence of the dynamic spatial categorization (DYSC) task and locations of neural recording in parietal (PAR) and prefrontal (PFC) cortex. Each trial we presented a small circular sample stimulus, and a line serving as a boundary cue. We varied the order of presentation of these stimuli, using either a sample-boundary sequence, or a boundary-sample sequence. ***A.*** Categorizing stimuli according to the left/right (LR) rule under the sample-boundary sequence. Trials began with the presentation of a central gaze fixation target (gaze fixation was required throughout the trial until the response was made). A sample stimulus was presented at one of 8 or 12 randomly selected locations for 400 ms, followed by a 900 ms delay period (delay 1). The LR rule was instructed by the presentation of the boundary cue in a vertical orientation for 400 ms, followed by a 900 ms delay period (delay 2). The monkey was required to determine the spatial category of the sample position stored in working memory by evaluating its spatial relationship to the boundary cue. The sample on this trial belongs to the spatial category 'right'. After delay 2, two choice stimuli were sequentially presented for 700 ms each in randomized order, one in the opposite spatial category as the sample (choice 1; 'left') and one in the same spatial category (choice 2; 'right'). The DYSC task is a delayed category match-to-sample design. The monkey

was rewarded (with a drop of juice) if it pressed the response key during the period of time that the matching choice was visible (located in the same spatial category as the sample; choice 2 on this trial). ***B.*** Categorizing stimuli according to the above/below (AB) rule under the sample-boundary sequence. The boundary cue is presented in a horizontal orientation instructing the AB rule on this trial, and choice 1 matches the spatial category of the sample ('above'). ***C.*** Categorizing stimuli according the LR rule under the boundary-sample sequence. In this case, the sample stimulus is assigned to a category based on a boundary cue stored in working memory. ***D, E.*** Locations of neural recordings in parietal and prefrontal cortex of Monkeys 1 and 2 relative to positions of the principal sulcus ('PS'), central sulcus ('CS') and intra-parietal sulcus ('IPS') as reconstructed from structural MRI images. The perspective is a top-down view of the left cerebral hemisphere. Anterior ('Ant'), posterior ('Post'), medial ('Med') and lateral ('Lat') directions are as indicated by the arrows. Larger open circles indicate inner diameter of recording chambers over parietal and prefrontal cortex. Smaller filled circles within each cortical area indicate regions sampled during neural recording. ***F.*** Eight position sample array. ***G.*** Twelve position sample array.

Task The DYSC task (Fig. 20) required monkeys to determine whether a small circle (the sample stimulus) was located to one side of a line (the boundary cue) or the other, and to press a response key when a choice stimulus appeared on the same side of the boundary cue later in the trial. The boundary cue (Fig. 20a, 'boundary cue') constituted a category boundary that divided space into two regions, each containing a set of points grouped into a category by virtue of bearing the same spatial relationship to the boundary. (For example when the boundary cue was vertical, all points to the left of the boundary constituted one spatial category; all points to the right another.)

In each trial, we presented first a sample stimulus followed after a delay by the category boundary (Fig. 20a, b; sample-boundary sequence), or we presented the two stimuli in reverse order, first the category boundary followed by the sample stimulus (Fig. 20c, boundary-sample sequence). Presenting the two stimuli at different times

separated by an intervening delay ensured that the two stimuli were never simultaneously visible, and could not therefore jointly influence the receptive fields of visual neurons. Sample stimuli were small circles (0.25-0.5° diameter; either white or yellow), and boundary cues were blue lines (13 or 21.5° in length), back projected by an LCD projector (NEC MT820) onto an otherwise dark translucent screen located 65 cm in front of the monkey. The location of the sample was selected randomly from a circular array of either 8 or 12 positions located on an imaginary circle at an eccentricity of 13 degrees from the fixation target (Fig. 20f, g; Monkey 1 performed the task with both 8 and 12 position arrays, Monkey 2 performed the task with the 8 position array). We located sample positions in the arrays so that they did not fall on the boundary cue when presented in either the vertical (Fig. 20a; 'boundary cue') or horizontal (Fig. 20b; 'boundary cue') orientations (to ensure that each position in the sample array had a defined spatial category for both boundary orientations).

Monkeys initiated the trial by directing their gaze straight ahead toward a fixation target (0.25° red circle) presented directly in front of them (Fig. 20a-c; 'fix'). If the monkey's gaze deviated by more than 2.5 degrees from the fixation target at any point during the trial, the trial was terminated without reward. After 500 ms of central fixation, we presented the first visual stimulus in the sequence, either the sample stimulus (Fig. 20a, b; 'sample') or the boundary cue (Fig. 20c; 'boundary cue'), for 400 ms (For a portion of the boundary-sample data, we presented the boundary cue for 1000 ms). We presented the boundary cue either in a vertical (Fig. 20a, c) or horizontal (Fig. 20b)

orientation for each trial, always centered on and passing through the gaze fixation target. After the disappearance of the first stimulus in the sequence, a delay period of typically 900 ms (but sometimes 400 ms) followed (Fig. 20a-c; 'delay 1'). At the end of delay 1, the second stimulus was presented for 400 ms – either the boundary cue (Fig. 20a, b) or the sample stimulus (Fig. 20c), depending on the stimulus sequence employed. An important aspect of the task design was that the rule-dependent category of the sample stimulus was only defined after the second stimulus was presented in either stimulus sequence (boundary followed by sample or the reverse). After the offset of the second stimulus, a second delay period followed (Fig. 20a-c; 'delay 2'; typically 900 ms but sometimes 400 ms). At the end of delay 2, we presented two choice stimuli (small circles), one at a time for 700 ms each in random sequence. One choice was always a match (located on the same side of the boundary cue and therefore belonging to the same spatial category as the sample), and the other choice was always a nonmatch (located on the other side of the boundary cue and belonging to the opposite spatial category). If the monkeys pressed the response key during the period of time that that matching choice stimulus was visible (Fig. 20a-c; 'response'; boxes with thicker outlines), we delivered a liquid reward (0.1 – 0.2 ml sweetened water). We terminated the trial whenever the monkey pressed the response key, whether correct or not (if the monkey responded during the first choice period we did not present the second choice; as in Fig. 20b). The direction of the required motor response did not vary over trials, so that neural signals varying with the spatial features of the task were not likely to reflect motor plans of varying direction. Monkeys performed sets of trials consisting of 12 repetitions of each

sample position under each of the two boundary orientations (192 trials for 8 position arrays).

On trials in which the boundary cue was vertical (Fig. 20a, c; 'boundary cue'), it instructed the LR (left/right) categorization rule, and monkeys were required to determine whether the sample stimulus was located to the left or right of the category boundary. On trials in which the category boundary was horizontal (Fig. 20b; 'boundary cue'), it instructed the AB (above/below) rule, and monkeys were required to determine whether the sample stimulus was located above or below the category boundary. Changing the boundary orientation over trials therefore required reclassifying the circular stimulus array into orthogonal sets of spatial categories. The boundary cue served to instruct a categorization rule in the sense that its orientation dictated which of two independently varying dimensions of the sample stimulus (horizontal or vertical position) was relevant to category membership and which grouping criterion to apply to that dimension to determine the spatial category. When the boundary cue was vertical, the position of the sample along the horizontal axis was relevant to category membership and its vertical position was irrelevant. When the boundary cue was horizontal, the converse was true. We varied boundary orientation either randomly over trials, or in blocks of random length (between 7 and 12 trials).

For the boundary-sample data in Monkey 1, we selected match and nonmatch choice positions from the same stimulus array used to position samples (Fig. 20f, g), with

the constraint that the two choices on each trial were located at equal distances from the active boundary. For the remaining data (including the sample-boundary data in Monkey 1, and both the boundary-sample and sample-boundary data in Monkey 2), we more fully randomized the position of choice stimuli, selecting random positions for match and nonmatch choices on a circle centered on the fixation target (at an eccentricity of 13°). Specifically, on LR rule trials, when the sample was in the left category, we placed the matching choice in the left category at a random angle between 120° and 240°. We defined 0° as the direction straight to the right of the fixation target, and incremented angle counterclockwise, so that the vertical category boundary corresponded to angles 90° (up) and 270° (down), and the range from 120° to 240° fell entirely to the left of the boundary. On LR rule trials when the sample was in the right spatial category, we similarly placed the matching choice in the right category at a random angle and between 300° and 60° (a range to the right of the boundary). On AB rule trials, the boundary corresponded to the angles 0° (or 360°) and 180°. On AB rule trials when the sample was in the above category, we placed the matching choice at a random angle between 30° and 150° (a range falling above the boundary), and when the sample was in the below category, we placed the matching choice at a random angle between 210° and 330° (a range falling below the boundary). Once the angle of the matching choice was selected within the ranges specified, we positioned the nonmatching choice on the opposite side of the boundary at one of the two locations (selected at random) on the circle of 13° eccentricity from the fixation target that were located at the same distance from the category boundary as the match (effectively, the position of the nonmatch was the match

position on each trial reflected about either the active category boundary or both category boundaries). This meant that each randomly selected match position was associated with two nonmatch positions (with equal probability), and vice versa, such that the exact position of one choice stimulus could not be predicted with certainty based on the position of the other. Randomizing choice position in this manner helped to ensure that monkeys had to categorize space in order to select the correct choice and therefore respond accurately in the task. (They could not, for example, readily utilize a look-up table to select the correct choice based on particular combinations of sample and choice positions.) Randomizing choice position further ensured that the position of the matching choice was unpredictable from trial to trial. This helped to dissociate neural signals reflecting the category of the sample (based on its spatial relationship to the boundary cue), from neural signals representing the anticipated position of the matching stimulus that the monkey planned to select (as this position could not be anticipated).

Neural Recording We recorded neural activity in prefrontal and posterior parietal cortex simultaneously using two 16 electrode Eckhorn microdrives (Thomas Recording, GMbH; Giessen, Germany). Electrodes were 70 μm o.d. glass coated platinum iridium fibers (impedance 1-2 MΩ). Each electrode was advanced independently under computer control. The electrical signal from each electrode was amplified (5000X) and filtered (between 0.5 Hz and 5 KHz). The waveforms of the action potentials of individual neurons were isolated on-line by two people using a combination of time-amplitude window discriminators (Bak electronics, Mount Airy, MD) and waveform discriminators

(Alpha Omega Engineering, Nazareth, Israel). The timing of spike occurrence was stored to disk with 40 μs resolution (DAP 5200a Data Acquisition Processor, Microstar Laboratories, Bellevue, WA). We typically isolated the action potentials of 15-25 neurons simultaneously in each cortical area.

Data Analysis

We analyzed the activity of single parietal and prefrontal neurons using analysis of variance (ANOVA) and covariance (ANCOVA) as well as linear regression in order to evaluate the influence of sample position, boundary cue orientation, and spatial category on the activity of single neurons. We also applied a pattern classification analysis to patterns of activity in neuronal populations and ensembles (groups of simultaneously recorded neurons) in order to decode the value of the same variables, and further quantify their strength of representation by neural activity. We applied both single neuron and decoding analysis to firing rates measured in successive time bins throughout the trial, to measure the variation in neuronal representation of task variables over time. We compared the results of these analyses when applied to neural activity recorded in parietal and prefrontal cortex, to compare the time course and strength of neural representation of sample position, boundary orientation and spatial category across the two cortical areas. We wanted to quantify the strength and timing with which the boundary orientation or categorization rule modulated category signals in the two cortical areas, as this captured the essential executive control operation required by the DYSC task (the rule-dependent neural representation of category). We also computed a category-selectivity index

analogous to that developed by Freedman, Miller and colleagues (Freedman, Riesenhuber et al. 2001; Freedman, Riesenhuber et al. 2002; Roy, Riesenhuber et al. 2010) to provide a quantification of category selectivity that was comparable with that published in prior reports.

ANOVA/ ANCOVA design We presented the sample and boundary cue in sequence during each trial of the DYSC task. This helped to parse influences of sample position, rule, and spatial category on neural activity. Except where noted, we utilized an alpha level of 0.05 to evaluate the significance of statistical tests. For neural data collected using the sample-boundary sequence, we evaluated the influence of sample position on firing rate averaged over the sample and delay 1 periods (Fig. 20a, b) using a one-way ANOVA with sample position (8 or 12 levels, depending on the sample array used) as the single factor. We performed a separate analysis on the sample-boundary data to evaluate the influence of boundary orientation and spatial category on firing rate averaged over the boundary cue and delay 2 periods (Fig. 20a, b), when rule-dependent category was explicitly defined by this stimulus sequence. In this analysis, we applied a 3-way ANCOVA, in which the factors were boundary orientation (LR or AB), horizontal sample category (left or right), and vertical sample category (above or below), defining sample category on each trial (regardless of boundary orientation) on the basis of whether the sample was located in the left, right, above, or below halves of the circular sample array. We included the mean firing rate on each trial averaged over the preceding sample and delay 1 period (Fig. 20a, b) as a covariate in the analysis. This allowed us to evaluate

the influence of boundary orientation and spatial category on firing rate apart from, and in addition to, any potential influence of sample position on activity carrying over from the sample and delay 1 period.

For neural data collected using the boundary-sample sequence, we evaluated the influence of boundary orientation on firing rate averaged over the boundary cue and delay 1 periods (Fig. 20c) using a one-way ANOVA with boundary orientation (LR or AB) as the single factor. We applied a separate analysis to quantify the influence of sample spatial category on neuronal firing rate averaged over the sample and delay 2 periods (Fig. 1c). Because sample position and spatial category were defined at the same moment by the presentation of the same visual stimulus, the analysis dissociating neural signals reflecting the position of the sample and its spatial category was necessarily more complex for the boundary-sample data. We performed two separate ANOVAs in this case, one testing neurons for significant horizontal category preference, the other testing neurons for a significant vertical category preference. Each analysis was a 3-way design in which the three factors were rule, category, and within-category position. In the analysis testing horizontal category preference, the category factor had two levels (left and right), and the within-category position factor had four or six levels corresponding to the vertical position of the sample in the 8 or 12 sample position arrays (Fig. 1f, g). In the analysis testing vertical category preference, the category factor had two levels (above or below), and the within-category position factor had four or six levels, reflecting the horizontal position of the sample, depending on the sample array used (Fig. 1g, h).

The rationale of each test was that neurons with legitimate category selectivity would exhibit activity that varied significantly across spatial categories, but not as a function of sample position within category. We therefore identified category-selective neurons as those in which firing rate varied significantly as a function of the category and not the within-category position factor for either the horizontal or vertical analysis (using an alpha level of 0.025 to evaluate the significance of each factor, employing the Bonferroni correction for multiple tests). We identified position-selective neurons as those in which firing rate varied as a function of within-category position in either analysis and that did not meet the criteria for either horizontal or vertical category-selectivity as defined above.

Decoding Analysis We applied the pattern classification analysis to the DYSC-rotate data in order to decode sample position, boundary orientation and spatial category from patterns of neuronal population activity, applying the analysis to neural activity patterns measured in successive 50 ms time bins to produce a time course of decoding accuracy over the trial. By measuring the proportion of trials that the decoding analysis predicted position, rule and category correctly at each time point, we were able to obtain a measure of fluctuation in the strength of neural population signals coding each task variable in prefrontal and parietal cortex. We selected neurons for inclusion in the neural populations used for the decoding based on the results of the ANOVA/ANCOVA. We ranked the cells based on the p-value associated with the behavioral variable that we were decoding. We then selected the most significant 70 neurons (lowest p-values), the most significant 200 neurons, or all significant neurons ($p<0.05$) to perform the decoding.

This approach allowed us to compare decoding accuracy while keeping the number of neurons constant across cortical areas as the number of neurons included in the analysis can influence decoding accuracy.

We performed the decoding in time resolved manner. We measured firing rates in successive 50ms time bins, and then at each time step included a series of consecutive 3-5 firing rate measurements for each neuron (spanning a period of 150-250 ms) into the vector of firing rates representing the population activity pattern. The reason we represented the activity of each neuron as a short series of successive firing rate measurements was that this improved decoding accuracy. For the decoding, we employed 5-fold cross validation. We divided the trials into 5 subgroups, and used 4/5 of trials as the training group, while the other 1/5 of trials provided the test data. We used neural activity on training trials to define the parameters of the classification functions, and used these functions to classify test trials based on patterns of population activity. For example, when we were decoding horizontal categories- LEFT/RIGHT, we divided training trials into two groups based on the horizontal category of the sample, and then computed the mean population activity vector for LEFT and RIGHT trials separately. To quantify the variance over trials of the population rate vector within spatial category, we computed the pooled within category covariance matrix. We then modeled the population code for each category as a multivariate normal distribution, with mean equal to the mean population rate vector computed for training trials of that category and variance quantified by the pooled within-category covariance matrix. To classify each

test trial, we computed the Euclidean distance between the population activity vector on that trial and each of the two mean vectors for LEFT and RIGHT in the training data, and converted these distances to posterior probabilities (using the pooled covariance matrix) based on the assumption that the distribution of the data in each group was multivariate normal (we used the *classify* function in the Matlab Statistics toolbox to compute the posterior probabilities). We then assigned the test trial to the group (LEFT or RIGHT) associated with the higher posterior probability, and accumulated the results over trials (within each time bin) to determine how often the analysis was correct. When decoding horizontal spatial category, we only used LR rule trials (as these were the trials in which horizontal categories were explicitly defined). Likewise, when decoding vertical categories, we only used AB rule trials.

Decoding rule – dependent category on error trials To evaluate whether neural signals coding rule-dependent category related to successful performance, we compared the accuracy of decoding category on the correct and error trials. We restricted the error analysis to neurons that were recorded during experiments in which the monkey committed at least 30 errors. Thus we calculated the decoding accuracy as a percentage of 30 error trials total. We ranked neurons according to the p-value associated with the interaction between rule and category, and selected the 50 most significant neurons to include in population. To increase the sensitivity of the analysis, we used a single time bin to analyze population activity starting from the onset of the second stimulus to the onset of the first choice stimulus. We decoded rule-dependent spatial category using four

groups: left, right, above, below, based jointly on sample position and boundary orientation. After training the classifier using activity on correct trials, we decoded rule-dependent spatial category on error trials. This allowed us to evaluate neural representation on error trials based upon the relation between patterns of activity and spatial categories observed during correct performance. We evaluated differences in decoding accuracy on correct and error trials using the z-test of proportions.

Category Selectivity Index We calculated a category selectivity index for neurons that had significant category signal found using ANOVA. The index for each neuron contrasted the average difference in firing rate associated with pairs of sample locations that belonged within the same spatial category against the average difference in firing rate for pairs of sample locations that belonged to different categories. To compute the within category difference (WCD), we computed the mean difference in firing rate for all pairs of sample positions oriented parallel to the category boundary so that both sample locations fell within the same category. To compute the between category difference (BCD), we repeated this using all pairs of sample positions oriented orthogonal to the category boundary so that the two positions fell in different categories. We defined the category-selectivity index for each neuron as the ratio of difference between the BCD and WCD divided by the sum of the BCD and WCD. We computed WCD and BCD values for neurons with left or right preferences separately on LR rule trials, when the rule was compatible with their category preference, and for AB rule trials, when the rule was incompatible with their category preference. We did the converse for neurons with above

or below category preferences. These constraints ensured that the number of sample pairs included, the distances between the samples in each pair, and the distance from the samples to the category boundary, were balanced across the BCD and WCD calculations.

Sliding Window Linear Regression Analysis We further evaluated the time course of population representation using a sliding window linear regression analysis. We measured the firing rate of each neuron in a 200 ms window advanced in 20 ms increments through the trial. At each time step, we fit firing rates to the following linear model:

$$F = \beta_0 + \beta_1 R + \beta_2 H + \beta_3 V + \beta_4 RH + \beta_5 RV + \varepsilon$$

Where F is firing rate, R is the categorization rule signified by the orientation of the boundary cue (LR or AB), H is the horizontal category of the sample (left or right), and V is the vertical category of the sample (above or below). The RH and RV terms capture the interaction between the categorization rule and horizontal and vertical categories, respectively. We used the *regstats* function of the Matlab statistical toolbox to estimate the regression coefficients β_0- β_5 at each time step of the analysis. The analysis produced a succession of regression coefficients for each neuron quantifying the magnitude of the relation between firing rate and a given behavioral variable as the rate window was advanced through the trial.

RESULTS

Behavioral performance The DYSC-rotate task had two variants based on the order in which stimuli were presented. In the boundary-sample sequence, the boundary cue was shown first in the trial and monkeys had to store the orientation of the boundary in working memory, so that it could be used when the sample appeared later to calculate the appropriate category of the sample based on its spatial relationship to the boundary. In sample-boundary sequence, the sample was shown first followed by the boundary cue. In this case, monkeys had to store the position of the sample in working memory and later use it to calculate the category when the boundary cue appeared. Staggering the presentation of the two stimuli and reversing their order in different experiments helped to dissociate neural signals reflecting the visual properties of the sample and boundary cue from other signals reflecting the spatial category of the sample. The spatial category of the sample was a cognitive variable jointly defined by both the sample and boundary cue but abstracted from their visual attributes when considered individually. The DYSC-rotate task required monkeys to flexibly parse the circular array into two orthogonal pairs of spatial categories – LEFT/RIGHT and ABOVE/BELOW, based on the changing orientation of the boundary cue. The orientation of the boundary cue therefore constituted a rule controlling how positions were allocated to categories.

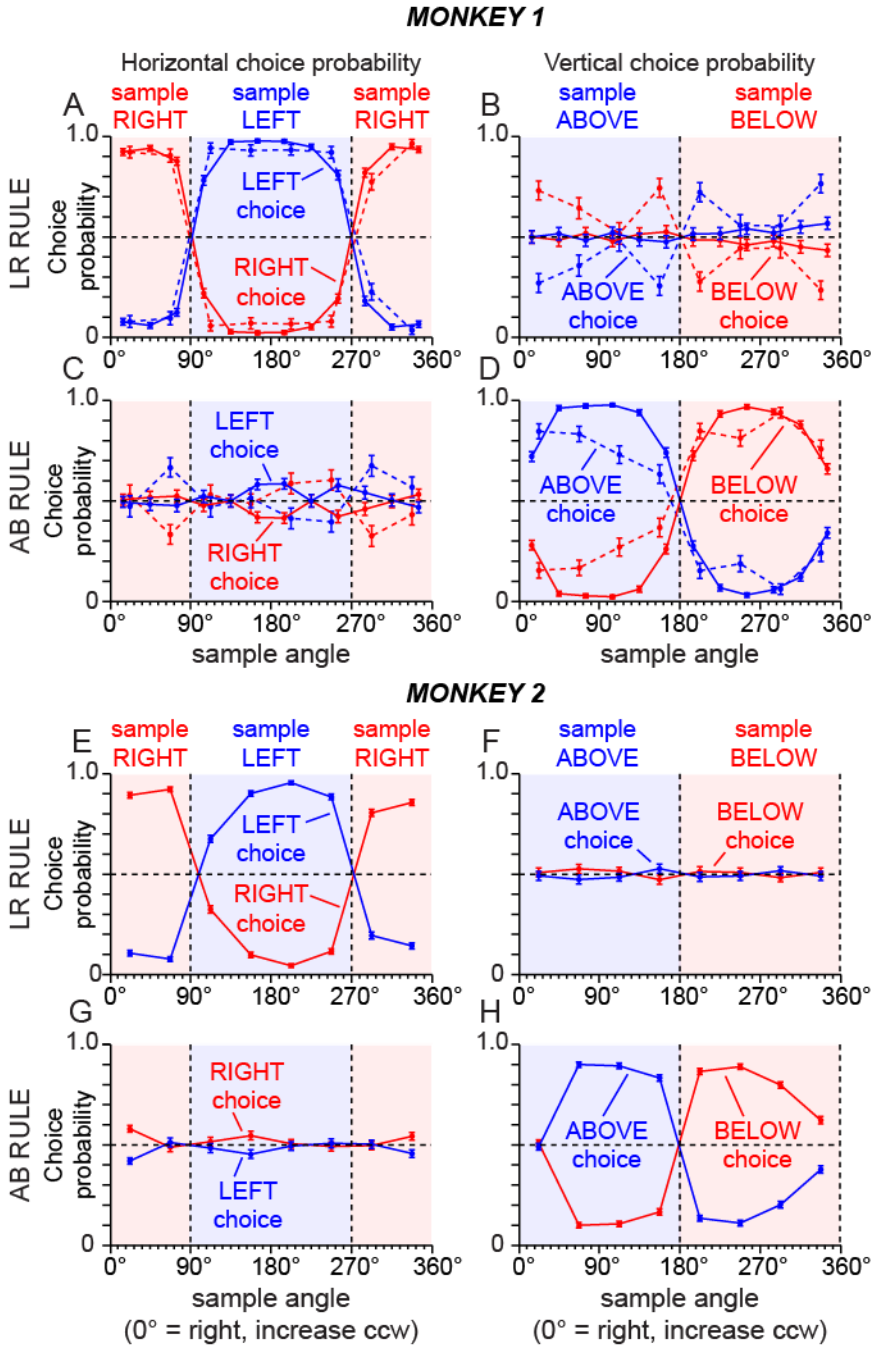

Figure 21. Behavioral performance of Monkeys 1 (*A-D*) and 2 (*E-H*) in the DYSC-rotate task. The horizontal axis of each plot indicates the position of the sample stimulus, in degrees angle counterclockwise from the direction to the right of the fixation target (defined as 0°; sample eccentricity was fixed at 13°). Spans of sample angle corresponding to the right and left halves of the circular sample array are indicated by red and blue shading, respectively (*A, C, E, G*). Spans of sample angle corresponding to the above and below halves of the circular sample array are indicated by blue and red

shading, respectively (*B, D, F, H*). Lines and symbols indicate for each sample position the probability (proportion of trials) that monkeys selected choices that were located in the right (red lines) or left (blue lines), above (blue lines) or below (red lines) spatial categories. Vertical black dashed lines indicate angles corresponding to the category boundary under a given rule. ***A, B.*** Performance of Monkey 1 under the LR rule, when the boundary cue was vertical. Solid lines indicated performance on trials using the boundary-sample sequence, dashed lines performance on trials using the sample-boundary sequence. ***A.*** Under the LR rule, Monkey 1 selected left choices with high probability (blue symbols and lines) when the sample category was left (blue shading), and right choices with high probability (red symbols and lines) when the sample category was right (red shading). ***B.*** Under the LR rule, the vertical category of the choices selected by Monkey 1 (above or below) did not relate systematically to the vertical category of the sample. ***C, D.*** Performance of Monkey 1 under the AB rule, when the boundary cue was horizontal. ***C.*** Under the AB rule, the horizontal category of the choices selected by Monkey 1 (left or right) did not relate systematically to the horizontal category of the sample. ***D.*** In contrast, under the AB rule, Monkey 1 selected above choices with high probability (blue symbols and lines) when the sample category was above (blue shading), and below choices with high probability (red symbols and lines) when the sample category was below (red shading). ***E – H.*** Corresponding data for Monkey 2.Red line refers to left rule (in LR rule) and below rule (in AB rule) whereas blue line represents right (in LR rule) and above rule (in AB rule). The sample array typically contained 8 but occasionally 12 positions.

Monkey 1 performed the DYSC-rotate task at 86.5% and 84.9% for early rule and late rule conditions respectively. Monkey 2 performed at 85% and 80% for early rule and late rule conditions respectively. Plots in Fig 21 represent the proportion of trials that monkeys selected the choice stimulus located in the LEFT/RIGHT, or ABOVE/BELOW spatial categories, based on the position of the sample and the categorization rule in force. Sample position is indicated by the angle of sample stimulus in polar coordinates counterclockwise from 0 degrees, defined as the horizontal direction to the right of the gaze fixation target. Under the LR rule, both monkeys selected choices in the left category with high probability on trials in which samples were located to the left of the category boundary (Fig. 21a, e; blue lines and regions). Left choice probability than abruptly dropped as the sample crossed the vertical LR category boundary into the right

spatial category (Fig. 12a, e; blue lines, red regions). Right choice probability was high when the sample was to the right of the category boundary (Fig. 12a, e; red lines and regions) and low when the sample was left of the category boundary (Fig. 12a, e; red lines, blue regions). Under the LR rule, the vertical category of the sample (above/below) had little systematic impact on vertical choice probability (21b, f), evidence that monkeys ignored the vertical coordinate of stimuli under the LR rule. Under the AB rule, monkeys exhibited the converse pattern, and selected choice stimuli based on the vertical category of the sample (Fig. 21d, h), and not the horizontal category of the sample (Fig. 21c, g). Categorization performance under the AB rule was less accurate than under the LR rule in general, but both monkeys clearly demonstrated rule-dependence in the way they categorized the stimuli.

We recorded the activity of 1016 neurons in parietal cortex and 977 neurons in prefrontal cortex of two monkeys performing the DYSC task. In monkey 1, we recorded the activity of 504 parietal neurons and 496 prefrontal neurons. Of these, 705 neurons were recorded using the boundary-sample stimulus sequence and 295 neurons using the sample-boundary stimulus sequence. In monkey 2, we recorded the activity of 512 parietal neurons and 481 prefrontal neurons using the sample–boundary stimulus sequence.

Neural signals coding sample position, rule, and spatial category in parietal and prefrontal cortex

The Venn diagram in Fig. 22 illustrates the numbers of neurons in parietal cortex and prefrontal cortex recorded using the boundary-sample and sample-boundary event sequences in which firing rate varied significantly ($p < 0.05$) as a function of sample position ('Pos'), boundary orientation ('Rule') and sample category ('Cat') in the analysis of variance and covariance (ANCOVA).

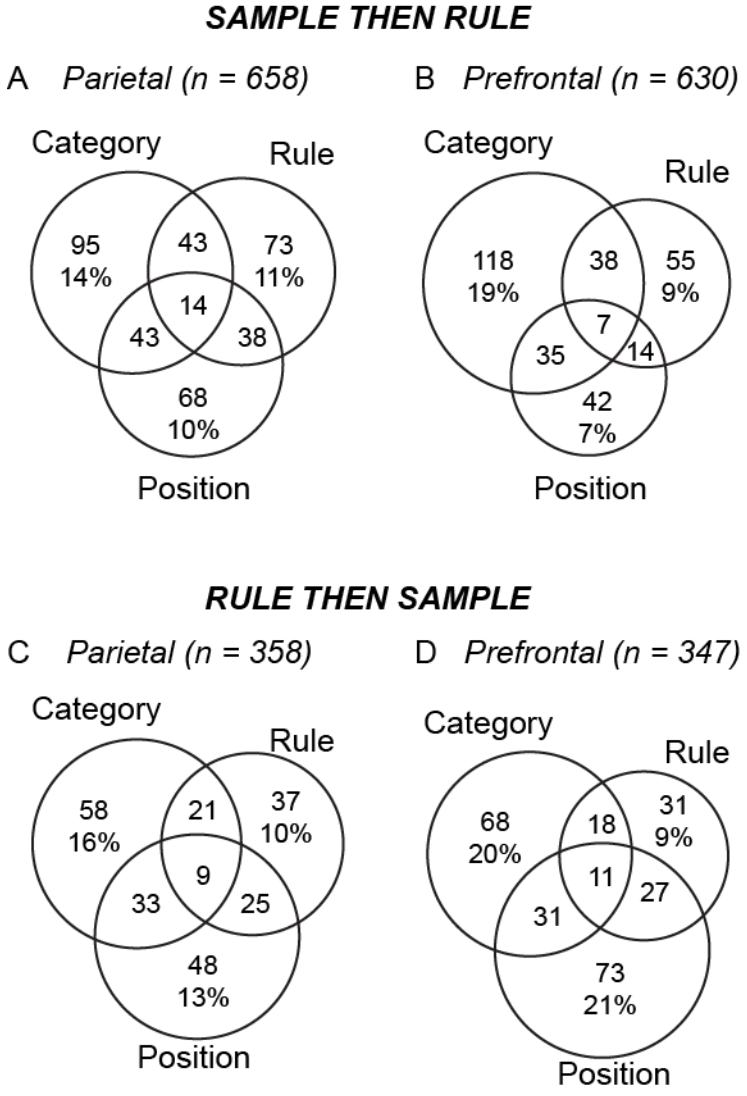

Figure 22. Proportions of neurons in(***A***) parietal cortex, and (***B***) prefrontal cortex having activity that varied significantly ($p < 0.05$) as a function of position, rule or category factors in the analyses of variance and covariance using the sample-boundary ('sample then rule') sequence. ***B, C.*** Corresponding data for neurons recorded using the boundary-sample ('rule then sample') sequence.

Neurons with activity relating significantly to position, rule and category were distributed between parietal and prefrontal cortex, when using both stimulus sequences. However, pure category neurons, in which firing rate related significantly to the category of the sample and not its position or the categorization rule were modestly concentrated in prefrontal cortex (Fig. 22). When we employed the sample-boundary stimulus sequence, the proportion of neurons coding only the spatial category of the sample stimulus and not the other factors was significantly greater ($\chi^2 = 8.8346$; $p = 0.003$) in prefrontal (Fig. 22b) than in parietal (Fig. 22a) cortex.

Figure 23 demonstrates the activity of representative neurons that code sample position, boundary position and spatial category both in parietal (Fig. 23a, c, e) and prefrontal cortex (Fig. 23b, d, f). We observed cells both in parietal and prefrontal cortex that coded the location of sample position. The trials were divided into preferred and non preferred trials based on preferred position of the cell. Fig. 23 C and D represents neurons that code the location of boundary. Interesting difference to note between the two cell is that the parietal neuron is a visual rule neuron meaning it is only active when boundary is visible on the screen whereas prefrontal neuron is boundary delay neuron and gets activated for a longer duration. It is suffice to note that we found both kind of cells in both areas.

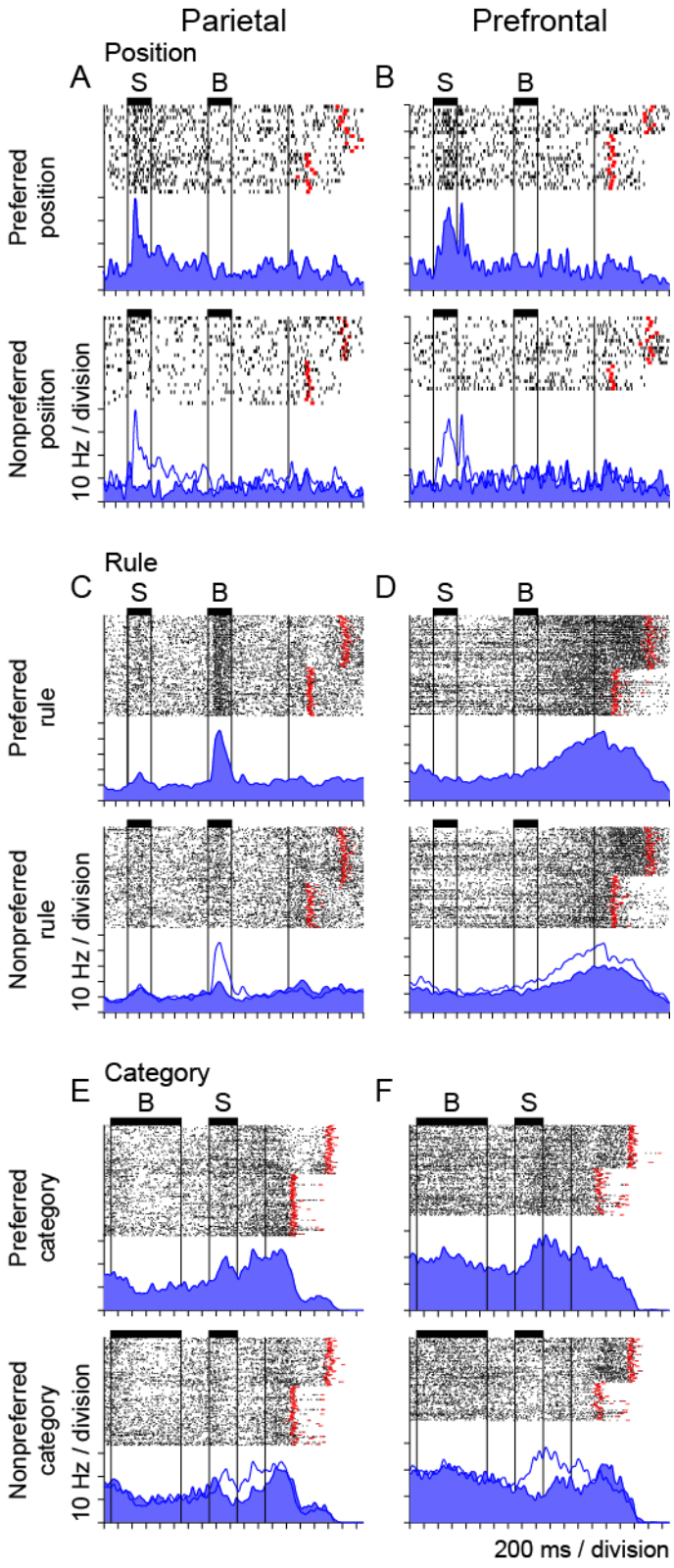
Parietal
Prefrontal
Position
A
B
S
B
Preferred position
Nonpreferred positon
10 Hz / division
Rule
C
D
Preferred rule
Nonpreferred rule
Category
E
F
Preferred category
Nonpreferred category
200 ms / division

Figure 23. Rasters and spike density functions illustrate the activity of single neurons in parietal cortex (***A, C, E***) and prefrontal cortex (***B, D, F***) varying significantly ($p < 0.05$) as a function of sample position (***A, B***), rule (or boundary orientation; ***C, D***), and spatial category (***E, F***). Neural activity associated with the preferred position, rule or category of each neuron is illustrated in the upper raster of each panel and activity on non-preferred trials is illustrated in the lower raster (the thin blue line in the lower rasters shows activity on preferred trials for comparison).

Finally, we did find neurons that coded the sample category both in parietal and prefrontal cortex (Fig. 23 e and f). To compare the strength and timing of physiological signals in parietal and prefrontal cortex coding sample position, rule, and spatial category, we applied a pattern classification analysis to decode these task variables from population activity patterns measured in successive 50 ms time bins on trials using the sample-boundary sequence. Population signals coding the position of the sample stimulus were either earlier (Fig. 24a) and/or stronger (Fig. 24b) in parietal cortex relative to prefrontal cortex in the two monkeys (Fig. 24; dots and thicker lines indicate a significant difference in the proportion of correctly decoded trials between cortical areas). Population signals coding the orientation of the category boundary (the rule) were likewise earlier and stronger within parietal cortex relative to prefrontal cortex in both monkeys (Fig. 24c, d). We obtained the reverse result in the relative strengths with which parietal and prefrontal cortex represented spatial category. Population signals reflecting the spatial category of the sample were stronger and earlier in prefrontal cortex relative to parietal cortex in both monkeys (Fig. 24e, f). These data are consistent with signals reflecting visual features of stimuli flowing in a forward direction from parietal to

prefrontal cortex, and signals reflecting abstract spatial categories in our task flowing in the reverse direction, from prefrontal to parietal cortex.

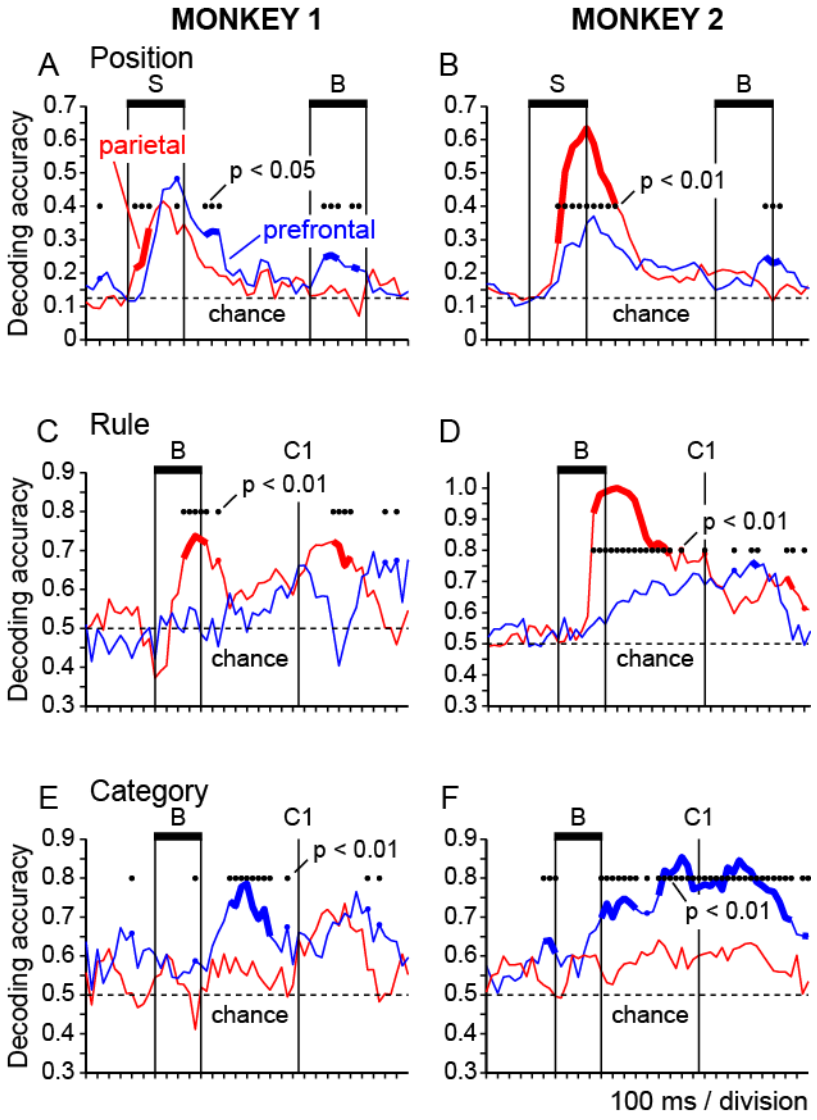

Figure 24. Plots of decoding accuracy over time indicate the proportion of trials in which sample position (A, B), rule (C, D) and spatial category (D, E) were accurately decoded from population activity patterns measured in successive 50 ms time bins in parietal cortex (red) and prefrontal cortex (blue). Data from monkey 1 (A, C, E) and monkey 2 (B, D, F) are plotted separately. Time bins in which the proportion of correctly decoded trials varied significantly between prefrontal and parietal cortex are indicated by filled circles and sections of the time courses plotted with a thicker line (z test of proportions, α level indicated in each panel).

Rule-dependent category signals in the prefrontal-parietal network

Modulation of neural signals coding spatial category as a function of a variable rule provided a neural correlate of executive control over a spatial cognitive process in the DYSC task. Figure 25 represents the rule-dependent category-selective activity of a single neuron in parietal (Fig. 25a, b) and prefrontal (Fig. 25c, d) cortex. The parietal neuron exhibited elevated activity relatively late in the trial (toward the time of the choice sequence) when the sample appeared in any of the positions in the ABOVE spatial category on AB rule trials when the category information carried by the activity of this neuron was relevant to the task, given the rule in force (Fig. 25b: positions 1-6), but not when the sample stimulus appeared in the same positions under the LR rule (Fig. 25a: positions 1-6) which rendered the category information carried by the cell irrelevant to the task. The activation of this neuron corresponded to the cognitive assignment of the sample to the spatial category ABOVE, which was the preferred spatial category for this neuron. The firing rate of this neuron observed for the six sample positions within the preferred category was similar, indicating that the level of activity remained essentially constant across different positions within a category. In addition, it is important to note that this neuron did not encode the category of the sample stimulus immediately when it appeared (black bars labeled 'S' indicate the timing and duration of the sample), rather the increase in discharge was evoked by the presentation of the boundary cue but only on trials in which the boundary cue, taken in combination with the position of the sample stored in working memory, jointly determined assignment of the sample stimulus to the

preferred spatial category (ABOVE for this neuron). Similarly, the prefrontal neuron (Fig. 25 c and d) showed preferential activity for above category but only under AB rule. The task stimulus contains only 8 peripheral targets in this case. We observed similar pattern of activity, in which the neuron was activated for one category but only under one rule. Thus, these neurons represent rule based category representation at a single cell level.

By the same token, the activity of these neurons did not depend solely on the orientation of the boundary cue either, or the neurons would have been active on all trials in which the boundary cue was horizontal, instructing the AB rule, and not just the subset of these trials associated with the upper category (Fig. 25b, d). Comparing the activity of the two neurons, the prefrontal neuron exhibited notably stronger category selective activity (Fig. 25d) relative to the parietal neuron (Fig. 25b), a pattern that is consistent with the differential representation of rule-dependent categories at the population level between the two cortical areas that we document below.

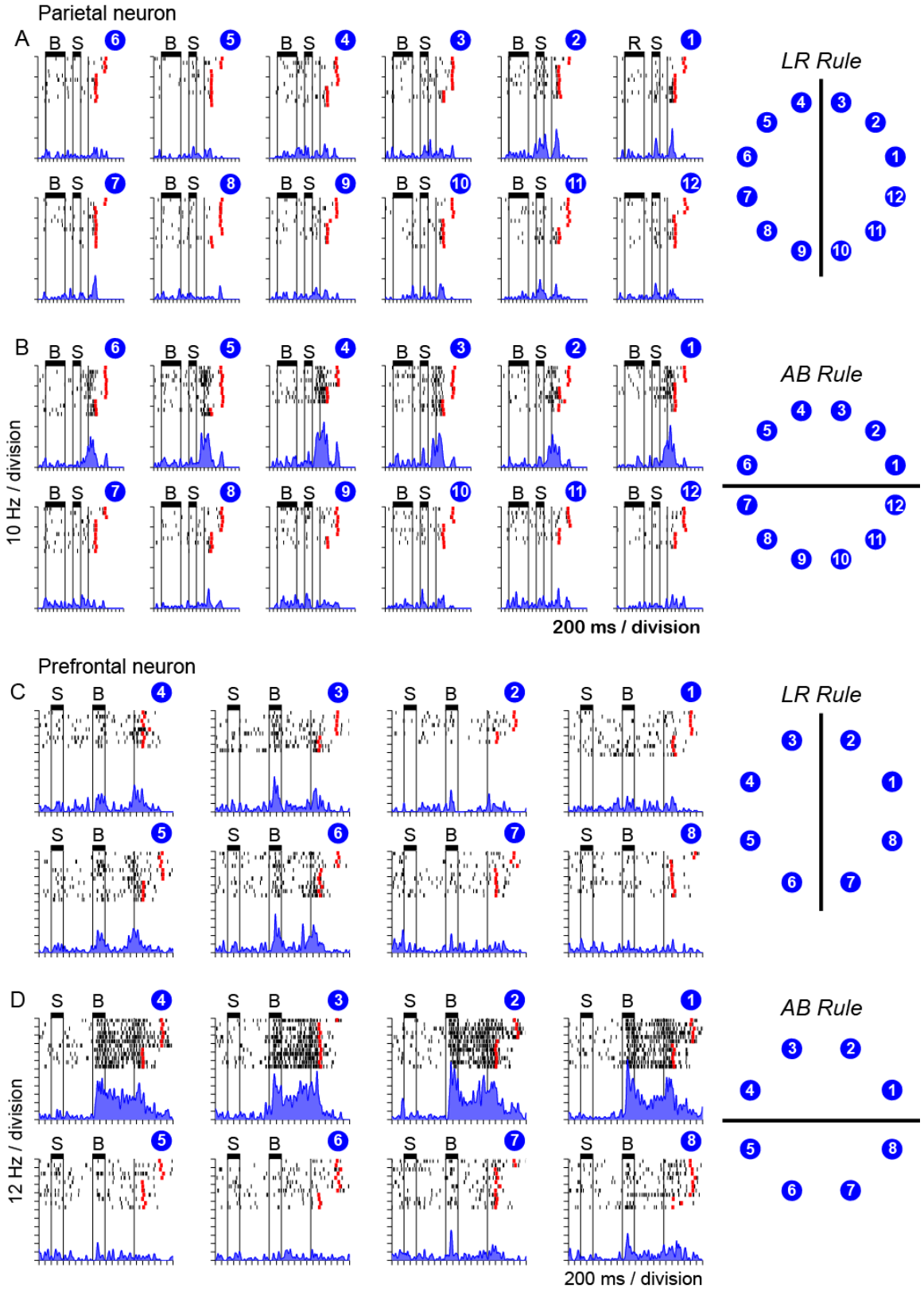
Parietal neuron
A
B
10 Hz / division
200 ms / division
LR Rule
AB Rule
Prefrontal neuron
C
D
12 Hz / division
200 ms / division
LR Rule
AB Rule

Figure 25. Rule-dependent category-selective activity in single neurons. ***A, B***. Activity of a neuron in parietal area 7a recorded using the boundary-sample sequence. Activity is plotted separately on trials performed under the LR rule (A) and the AB rule (B). Rasters are numbered according to the position of the sample stimulus presented each trial (diagram at right). This parietal neuron was activated when the sample was located in the 'above' spatial category under the AB rule (B) but not the LR rule (A). ***C, D.*** Activity of a neuron in prefrontal area 46 recorded using the sample-boundary sequence. This prefrontal neuron was also activated when the sample was located in the 'above' spatial category under the AB rule (D) but not the LR rule (C).

To examine the time course of population activity that reflected the rule-dependent spatial category of the sample in parietal and prefrontal cortex, we calculated the average population spike density functions (SDF) for groups of neurons in the two area using neurons with activity significantly ($p < 0.05$) influenced by the interaction between spatial category and rule in the ANCOVA. We constructed separate population activity functions for neurons coding the horizontal categories LEFT and RIGHT (Fig. 26a, c, e, f) and the vertical categories ABOVE and BELOW (Fig. 26b, d, g, h), on trials performed under the LR rule (left panel of each pair) and the AB rule (right panel of each pair). Neural populations in both parietal cortex (Fig. 26a, b, e, g) and prefrontal cortex (Fig. 26c, d, f, h) tended to only represent spatial categories that were relevant given the categorization rule in force, exhibiting the dual dependence on category and rule serving as a neurophysiological measure of executive control over a spatial cognitive process. For example, neurons that preferred the categories either LEFT or RIGHT were differentially activated to code these categories under the LR rule (Fig. 26a, c, e, f; left panel) but not the AB rule (Fig. 26a, c, e, f; right panel). Neurons preferring the categories ABOVE and BELOW exhibited similar rule-dependence, and were activated

differentially to code vertical categories during the AB rule and not the LR rule (Fig. 26b, d, g, h).

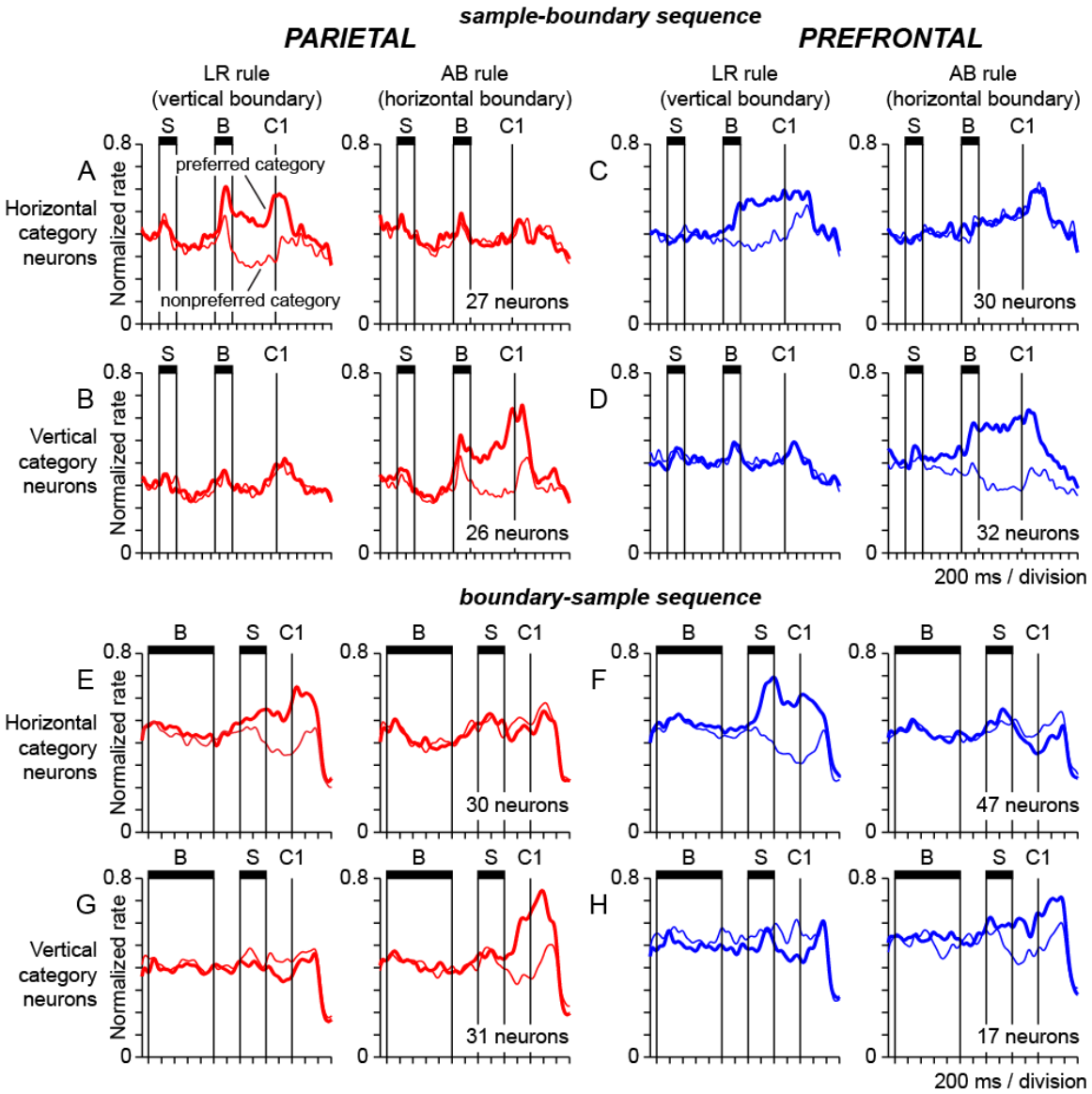

Figure 26. Average normalized population spike density functions (SDFs) (constructed using a Gaussian kernel of width $\sigma = 40$ ms) illustrate the activity of neurons coding rule-dependent category in parietal (red) and prefrontal (blue) cortex. SDFs plotted with a thicker line indicate population activity on preferred category trials; SDFs plotted with a thinner line indicate activity on non-preferred category trials. Neurons were included if their activity during the boundary and following delay periods related significantly ($p <$

0.05) to horizontal category and its interaction with rule, or vertical category and its interaction with rule. Neurons were excluded if their activity related to sample position ($p < 0.1$) during the sample or subsequent delay periods. Vertical lines delineate the sample period (black bar labeled 'S') and the boundary cue period (black bar labeled 'B'), and indicate the presentation of the first choice ('C1'). ***A, C.*** Population SDF illustrate the average activity of 27 parietal neurons (*A*) and 30 prefrontal neurons (*C*) that exhibited a horizontal category preference (left or right). Activity of the same neurons is plotted separately on LR rule trials (*A, C*; left plot of the pair in each panel) and AB rule trials (*A, C*; right plot of each pair). Neural data were collected using the sample-boundary sequence – data from both monkeys averaged. ***B, D.*** Population SDFs illustrate the average activity of 26 parietal neurons (*B*) and 32 prefrontal neurons (*D*) that exhibited a vertical category preference (above or below). Activity of the same neurons is plotted separately on LR rule trials (*B, D*; left plot of the pair in each panel) and AB rule trials (*B, D*; right plot of each pair). Neural data were collected using the sample-boundary sequence – data from both monkeys averaged. ***E - H.*** Corresponding average population spike density functions of neural activity recorded using the boundary-sample sequence (data from Monkey 1).

Moreover, population signals that carried rule-dependent category information emerged when the second stimulus in the trial sequence was shown, whether boundary cue or sample, and regardless of whether the sample-boundary sequence (Fig. 26a-d) or the boundary-sample sequence (Fig. 26e-h) was used. The presentation of the second stimulus in each sequence was the point in time at which sufficient information had been provided by the stimuli in the trial to compute the rule-dependent category of the sample.

Category Selectivity Index We calculated a category selectivity index for neurons with activity significantly influenced by either the horizontal or vertical category factor in the analysis of variance and covariance. The value of the category selectivity index varied from +1 to indicate pure category selectivity (neurons having a difference in mean firing rate between categories but not within categories), to -1 to indicate no category selectivity (difference in firing rate within categories but not between categories). We computed the

selectivity index separately using activity on two groups of trials, when the rule was compatible and incompatible given the category preference of each neuron (for example – if a neuron had a left spatial category preference, LR rule trials were compatible whereas AB rule trials incompatible). From Figure 27, we can see that when the rule was compatible, the distribution of the category index value was skewed to positive values in both parietal and prefrontal cortex, whereas when the rule was incompatible, the distribution of the category index value was more closely centered near zero for both parietal and prefrontal cortex. The degree of category-selectivity evaluated by this index was significantly modulated by the categorization rule in both cortical areas. The mean category index was significantly greater when the rule was compatible (Fig. 27a, b), than when it was incompatible (Fig. 27c, d), both in parietal cortex (t-test; 0.23 versus -0.05; $p < 0.0001$), and prefrontal cortex (t-test; 0.25 versus -0.03; $p < 0.0001$).

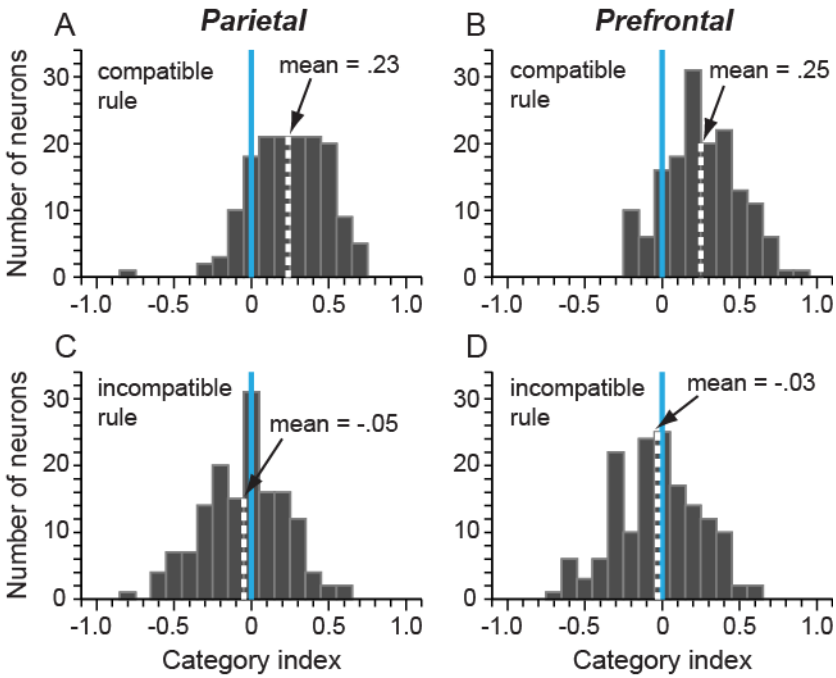

Figure 27. Distributions of the category index values of single neurons recorded using the sample-boundary sequence in parietal and prefrontal cortex. Neurons were included if their activity related significantly to either horizontal or vertical category ($p < 0.05$ by ANCOVA). The vertical white dashed line indicates the mean value of the category index in each distribution (blue line indicates an index value of 0). ***A.*** Distribution of category index values in parietal cortex on trials in which the rule was compatible with the category preference of each neuron (for example, LR trials for neurons with either left or right category preferences). ***B.*** Distribution of category index values in prefrontal cortex on compatible rule trials. ***C.*** Distribution of category index values in parietal cortex on incompatible rule trials (same neural sample as in A). ***D.*** Distribution of category index values in prefrontal cortex on incompatible rule trials (same neural sample as in B).

Evaluating rule-modulation of category signals: population decoding We applied pattern classification to further measure and characterize the rule modulation of category signals in the parietal-prefrontal network. To decode horizontal categories, we selected neurons based on the extent to which their activity was influenced by horizontal category in the analyses of variance and covariance, and decoded horizontal category in two separate analyses using the activity of these neurons on trials performed under the compatible (LR) and incompatible (AB) rules. We similarly decoded vertical categories on compatible (AB) and incompatible (LR) rule trials. Changes in decoding accuracy as a function of the rule marks a neurophysiological correlate of executive control in this task. Figure 26 shows that decoding accuracy obtained on compatible rule trials (red and blue lines) was markedly greater than decoding accuracy on incompatible rule trials (black lines) in parietal cortex (Fig. 28a, d, g) and prefrontal cortex (Fig. 28b, e, h). Decoding results are plotted separately based on the activity of the most significant 70 neurons (solid lines), the most significant 200 neurons (dashed lines), or all significant neurons (dotted lines).

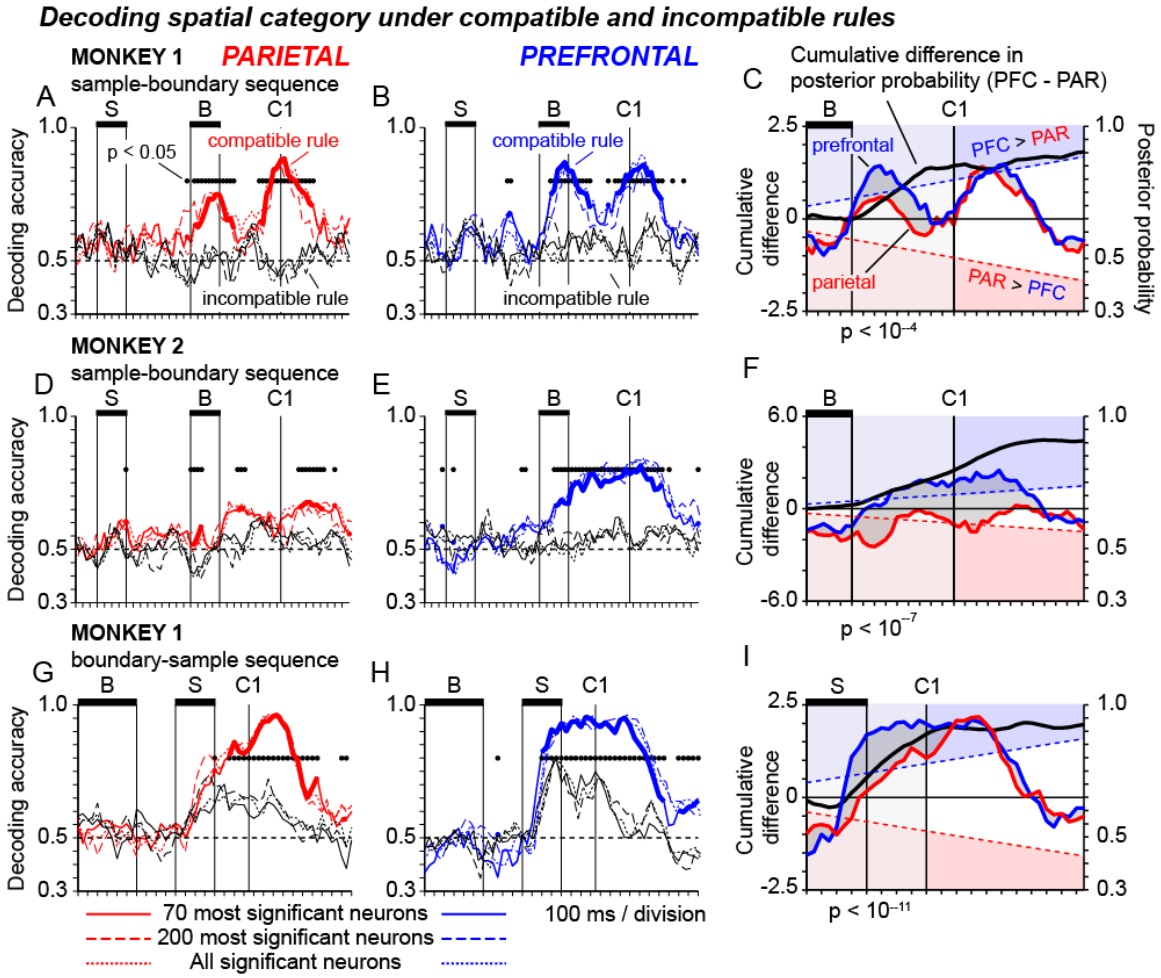

Figure 28. Decoding time courses illustrate fluctuation in the accuracy of decoding spatial category based on population activity patterns in parietal cortex (red) and prefrontal cortex (blue), measured in successive 50 ms time bins. Separate decoding time courses in each panel plot the results obtained using population activity on trials in which the rule was compatible with the category preference of each neuron (red and blue), and trials in which the rule was incompatible (black). Time bins in which the proportion of correctly decoded trials differed significantly ($p < 0.05$, z-test of proportions) between compatible and incompatible rule trials are indicated by black circles and corresponding thicker sections of the decoding time courses. Neurons in this analysis were ranked according to the p-value associated with the interaction between rule and category in the ANOVA/ANCOVA, and then varying numbers of the most significant neurons were selected to include in the populations used for the decoding. Decoding results indicate the accuracy obtained when based on the activity of the most significant 70 neurons (solid lines), the most significant 200 neurons (dashed lines), or all significant neurons (dotted lines) in each cortical area and monkey. ***A, B.*** Decoding accuracy obtained when using population activity in (*A*) parietal cortex, and *(B)* prefrontal cortex of Monkey 1 on sample-boundary trials. The difference between decoding accuracy on compatible and

incompatible rule trials measures the modulation of category signals by the rule. For the decoding analysis including all significant neurons, 55 neurons in parietal cortex and 58 neurons in prefrontal cortex contributed. ***C.*** Blue and red time courses indicate the mean posterior probability over time (averaged over trials) associated with the correct spatial category on compatible rule trials, based on neural activity in prefrontal and parietal cortex respectively (data from Monkey 1 on sample-boundary trials). The black time course illustrates the cumulative difference between the two time series. The diagonal dashed lines indicate the upper and lower confidence boundaries established by the sequential trials test. The cumulative difference function crosses the upper confidence boundary, indicating that the posterior probability is significantly larger in prefrontal cortex relative to parietal cortex ($p < 0.05$). ***D – F.*** Corresponding data from Monkey 2 on sample-boundary trials. For the decoding analysis including all significant neurons, 62 neurons in parietal cortex and 90 neurons in prefrontal cortex contributed. ***G – I.*** Corresponding data from Monkey 1 on boundary-sample trials. For the decoding analysis including all significant neurons, 104 neurons in parietal cortex and 120 neurons in prefrontal cortex contributed.

We performed the decoding analysis based on firing rates measured within 3 consecutive 50 ms time bins (constituting a 150-millisecond sliding window), which was advanced by 50-millisecond steps. We coded trials in which the decoding returned the correct category as 1, and trials in which the analysis returned the incorrect category as 0. We then applied the z-test of proportions to compare the proportion of hits and misses on compatible versus incompatible rule trials in the two cortical areas in each time bin (Fig. 28a, b, d, e, g, h; black dots in a horizontal row indicate time bins where accuracy differed significantly as a function of trial compatibility at $p < 0.05$). This analysis confirmed that category signals were modulated by rule compatibility in both prefrontal cortex (Fig. 28a, d, g) and parietal cortex (Fig. 28b, e, h). This provides evidence that the executive control of cognitive processing in the task was distributed and a network level process. The analysis also provided evidence that rule-modulated category signals were

somewhat stronger in prefrontal cortex (Fig. 26: blue) than in parietal cortex (Fig. 26: red).

To confirm that the modulation of category signals by the rule was significantly stronger in prefrontal cortex than parietal cortex we compared the time courses of posterior probabilities produced by the decoding analysis applied to neural activity in prefrontal cortex (Fig. 28c, f, i; blue lines) and parietal cortex (Fig. 28c, f, i; red lines), using the 70 most significant neurons in each cortical area, and restricting the data to that recorded on trials performed under the compatible rule. We computed the cumulative difference between the two time series (Fig. 28c, f, i; thick black time course) and tested the significance of this difference using the sequential trials test. The cumulative difference function crossed the upper confidence boundary established by the test indicating that posterior probabilities (which measure the certainty of category decoding) were significantly stronger in prefrontal cortex than parietal cortex (Fig. 28c, f, i), in both monkeys, and using both stimulus sequences.

We confirmed this result by also testing the significance of the difference in rule-modulated category signals across cortical areas using a two-tailed t-test contrasting the mean posterior probability associated with the correct rule-dependent category on compatible-rule trials in parietal and prefrontal cortex. The mean posterior probability observed in the interval from the onset of the second stimulus to the first choice (Fig. 28) was significantly greater in prefrontal cortex than parietal cortex in Monkey 1 on sample-

boundary trials (Fig. 28c; .62 versus .57, $p < 10^{-4}$), monkey 2 on sample-boundary trials (Fig. 28f; .59 versus .55, $p < 10^{-7}$), and Monkey 1 on boundary-sample trials (Fig. 28i; 0.74 versus 0.63; $p < 10^{-11}$).

We further detected a significant interaction between cortical area and rule compatibility on the posterior probability associated with the spatial category of the sample in a two-way ANOVA, with the rule exerting a significantly stronger effect on posterior probability when decoding was based on neural activity in prefrontal cortex in comparison to parietal cortex, both in monkey 1 ($F = 5.72$, $p < .05$), and monkey 2 ($F = 8.02$, $p < .05$).

Evaluating rule-modulation of category signals: regression analysis To provide an alternate measure of category signals and their modulation by the rule at the single neuron level, we measured the firing rate of every neuron (without exclusion) within a sliding 200ms window advanced in 20ms steps, and at each time step regressed the resulting rate of firing over trials onto category and its interaction with rule in a linear regression analysis (including the rule, horizontal category, vertical category, and their interactions as independent variables). This resulted in a time series of regression coefficients for each neuron measuring the strength of the relationship between activity and spatial category as well as the interaction between category and rule. Figure 29 illustrates the results of this analysis (each row illustrates the time series of regression coefficients obtained for a single neuron normalized to the peak coefficient for that

neuron, warmer colors indicate larger values). Single neurons carried a peak signal related to category or the interaction with rule for a short period of time (short spans of warm colors in each row). We arranged neurons in order of increasing time to the peak coefficient within the trial (Fig. 28, bottom to top in each panel). The diagonal bands of color provide a type of a recruitment curve showing the rate at which neurons were recruited in parietal cortex (Fig. 29a, b, e, f) and prefrontal cortex (Fig. 29c, d, g, h) to code spatial category and its interaction with rule (the white line in each panel shows the time course of recruitment in the companion cortical area for comparison). Pairs of plots illustrating recruitment in parietal and prefrontal cortex for a given variable include equal numbers of neurons in the two areas so the slopes of the lines of color indicating neuronal recruitment can be visually compared across areas. Neurons with activity reflecting spatial categories (Fig. 29a, c; e, g) and the interaction between rule and category (Fig. 29b, d, f, g) were recruited at a faster rate or at earlier times in the trial in prefrontal cortex (Fig. 29c, d, g, h) compared with parietal cortex (Fig. 29a, b, e, f), in both Monkey 1 (Fig. 29a–d) and Monkey 2 (Fig. 29e–h).

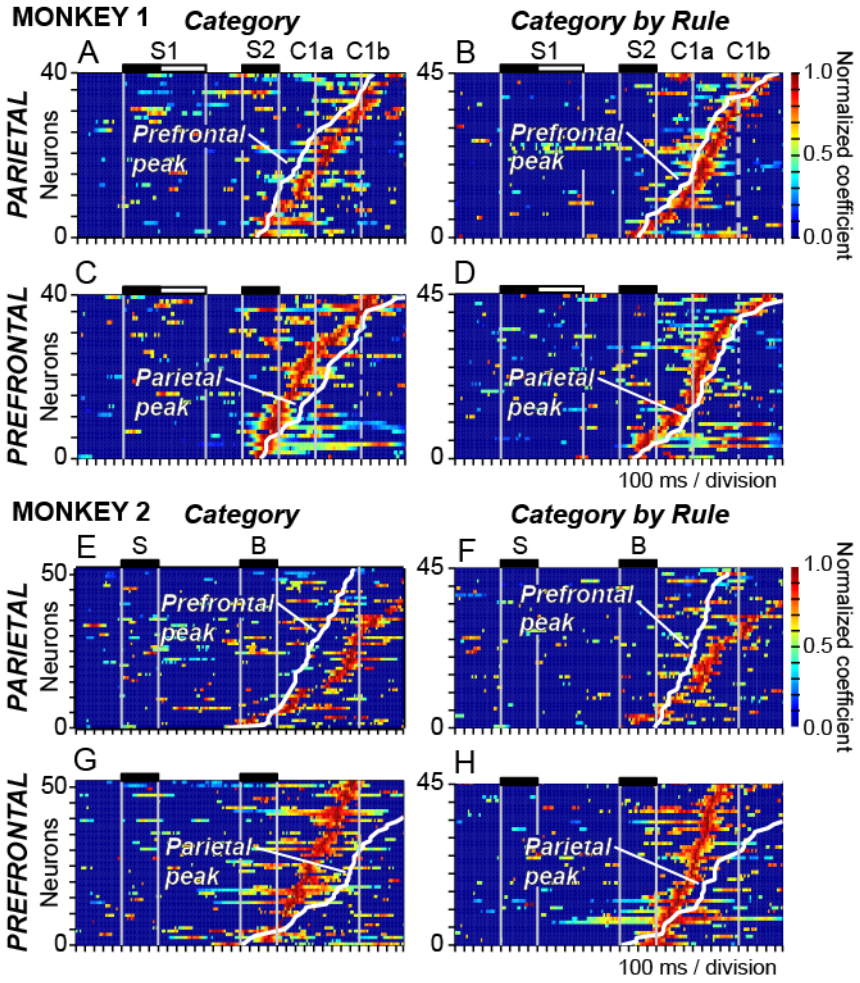

Figure 29. Results of an analysis regressing firing rates measured in a sliding 200 ms window, (advanced in 20 ms steps through the trial) onto category, and the interaction between category and rule. Rows in the color plots indicate the time series of regression coefficients obtained for single neurons normalized to the peak coefficient for each. Warmer colors indicated larger coefficients (scale at right). Coefficients at each time step were included in the time series only if they were significant ($p < 0.05$; dark blue areas indicates time steps where coefficients were not significant). Neurons (rows) are arranged according to the time of the peak coefficient, producing a representation of the rate at which neurons were recruited to represent category and the interaction between rule and category in parietal and prefrontal cortex. Vertical lines and black bars indicate timing of sample ('S') and boundary cue ('B') periods, and indicate the presentation of the first choice ('C1'). ***A-D.*** Time series of normalized regression coefficients obtained in Monkey 1 for category *(A, C)* and the interaction between category and rule *(B, D)* in parietal cortex *(A, B)* and prefrontal cortex *(C, D)*. Data from boundary-sample and sample-boundary trials are combined (and aligned to the onset of the second stimulus in either case, S2). ***G-L.*** Corresponding data from Monkey 2 (data from sample-boundary trials).

We also calculated the mean population regression coefficient for the interaction between rule and category in parietal and prefrontal cortex, and focused this analysis only on cells carrying the strongest signals (significant for the interaction between category and rule at $p<0.01$). We found in this analysis that the rising edge of the increase in the mean population coefficient for the interaction between rule and category lead in prefrontal cortex (Fig. 30ab; blue) and followed in parietal cortex (Fig. 30a, e; red) in both monkeys, and that the regression coefficient for the interaction was significantly larger in prefrontal cortex than parietal cortex by the sequential trials test in both monkeys (Fig. 30c, d). We performed a two-tailed t-test to confirm that the mean interaction coefficient in prefrontal cortex was significantly larger than the interaction coefficient in parietal cortex averaged over the interval between the onset of the second stimulus and the first choice both in Monkey 1 (mean parietal and prefrontal coefficients: 0.47 and 0.69 respectively, $p < 0.001$ two-tailed t-test) and in Monkey 2 (mean parietal and prefrontal coefficients: 0.49 and 0.99 respectively, $p < 10^{-9}$). The interaction coefficients are small because they are averaged over all neurons, most of which did not yield significant coefficients at any given time step (Fig. 29), contributing a value of 0 to the population average.

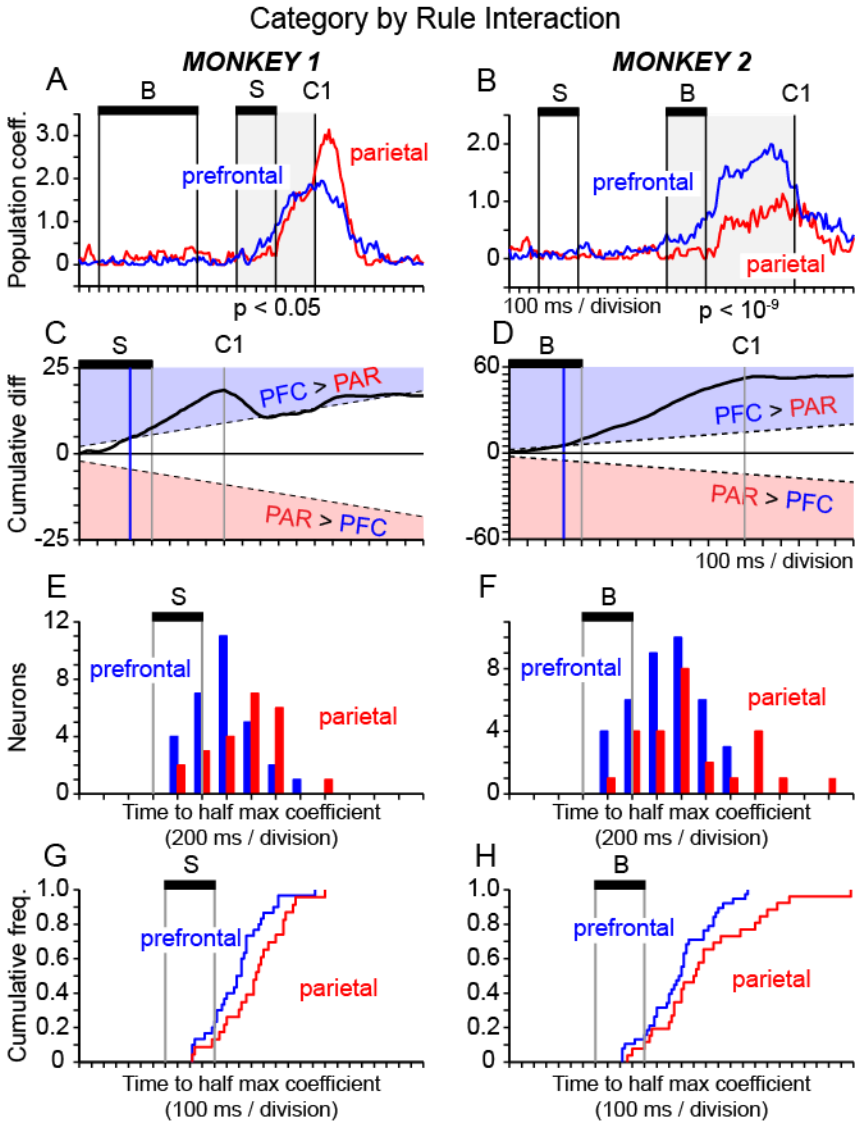

Figure 30. Comparison of the strength and timing of regression coefficients obtained for the interaction between rule and category in parietal and prefrontal neurons (data restricted to regression coefficients that were significant at $p < 0.01$). ***A, B.*** Plots illustrate variation over time in the mean population regression coefficient for the interaction in parietal cortex (red lines) and prefrontal cortex (blue lines) of Monkey 1 (*A*) and Monkey 2 (*B*). ***C, D.*** Results of sequential trials tests applied to population coefficient time series in parietal and prefrontal cortex in Monkey 1 (*C*) and Monkey 2 (*D*), starting at the onset of the second stimulus in the trial sequence. Thick black lines indicate the cumulative difference between population coefficient time series in the two cortical areas (prefrontal – parietal). The cumulative difference functions deflect upward and cross the upper confidence boundary (at the points indicated by vertical blue lines), indicating that the coefficient for the interaction between rule and category is significantly larger in prefrontal cortex ($p < 0.05$, minimum effect size of 0.8). ***E, F.*** Frequency distribution histograms of the time to the half-maximum regression coefficient for the interaction between rule and category in prefrontal cortex (blue bars) and parietal cortex (red bars) of Monkey 1 (*E*)

and Monkey 2 (*F*). ***G, H.*** Cumulative distributions of the time to the half-maximum regression coefficient for the interaction between rule and category in prefrontal cortex (blue lines) and parietal cortex (red lines) of Monkey 1 (*G*) and Monkey 2 (*H*).

To evaluate the significance of differences in the timing of this signal between cortical areas, we compared the mean time to the half maximum coefficient value in each neuron in the parietal and prefrontal populations. In Monkey 1, the time to the half maximum coefficient in prefrontal cortex (Fig. 30e, g; blue bars and line; mean = 582.7 ms after onset of the 2nd stimulus) was significantly earlier than the time to half maximum in parietal cortex (Fig. 30e, g; red bars and line; mean = 726.1 ms; $t = 2.05$, $p < 0.05$). Likewise in Monkey 2, the time to the half maximum coefficient in prefrontal cortex (Fig. 30f, h; blue bars and line; mean = 679.5 ms) was significantly earlier than the time to half maximum coefficient in parietal cortex (Fig. 30f, h; red bars and line; mean = 886.15 ms; $t = 2.33$, $p < 0.05$).

Together, the above data provide evidence that rule-modulation of category signals was stronger and earlier in prefrontal cortex relative to parietal cortex, suggesting that prefrontal cortex may initiate or direct the change in cortical representation associated with flexible cognitive processing in this task.

Relation of category representation to performance To determine whether the neural signal coding rule-dependent spatial category was related to behavioral performance, we compared the accuracy of decoding rule-dependent categories using neural activity that preceded the monkeys' correct (Fig. 31; black bars) and erroneous (Fig. 31; gray bars)

choices. The results of this analysis indicated that decoding accuracy was significantly reduced preceding an erroneous choice, approaching chance levels in both cortical areas as well as in both monkeys. Erroneous choices were therefore preceded by a failure of the prefrontal-parietal network to accurately encode the correct rule dependent spatial category of the sample, providing a link between this neural signal and behavioral performance.

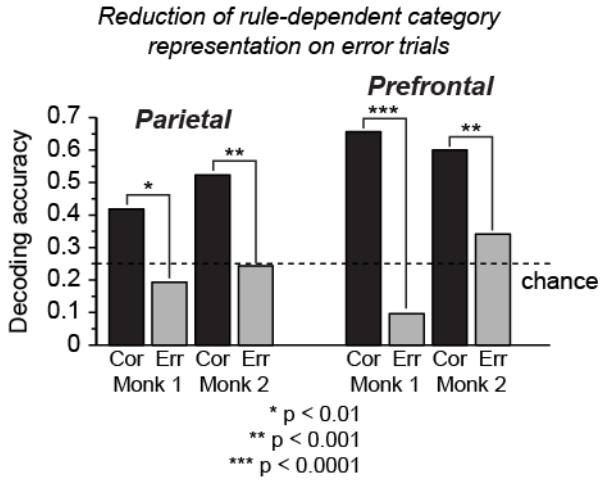

Figure 31. Relation of population decoding accuracy to behavioral performance. In this analysis, population activity was used to decode rule-dependent spatial category defined jointly by sample position and boundary orientation and coded as a categorical variable with four levels (left, right, above, and below). Decoding is based on the 50 neurons in each cortical area and monkey having activity most significantly related to the interaction between rule and category. Decoding accuracy was significantly reduced on error trials (gray) relative to correct trials (black), both in parietal cortex and prefrontal cortex in Monkeys 1 and 2 considered individually. The significance of the difference in decoding accuracy on correct and error trials, evaluated using the z-test of proportions, is as indicated by asterisks in the figure. Chance decoding (given 4 categories) is 25% correct.

Discussion

Executive control involves determining not only how to act, but how to think, or more specifically what cognitive process to execute to select the best action in a given context. Our objective in this experiment was to better understand how physiological signals in prefrontal cortical networks mediate executive control over cognitive processing. To approach that goal, we developed a task in which monkeys had to selectively implement competing cognitive processes in response to a rule cue. We then sought to isolate neural signals that reflected rule-dependent cognitive processing in the prefrontal-parietal network. The task we developed for this purpose required monkeys to flexibly reassign the same circular array of stimulus positions into different spatial categories according to different categorization rules or grouping criteria on a trial-by-trial basis. This task allowed us to isolate neural signals in the network that were selective for spatial categories and were also rule dependent, providing a cellular correlate of executive control over a cognitive process. By characterizing the timing and distribution of this physiological signal in the prefrontal-parietal network, we sought to understand further how the network implemented the computational flexibility the task required, and what the respective roles of parietal and prefrontal cortex in this computational flexibility might be.

We report several primary findings. First, the spatial variables defined by the DYSC task—spatial position, rule (boundary orientation), and spatial category—were each represented at the cellular level by the parallel activation of neurons with similar

physiological properties in parietal and prefrontal cortex. Neural representations related to task performance therefore were fundamentally distributed in this network. However, population decoding revealed that signals reflecting the spatial attributes of visual stimuli (sample position and boundary orientation) emerged earlier and/or were stronger in parietal cortex (area 7a) relative to prefrontal cortex (area 46), consistent with a feed forward of visually evoked signals in the network. Conversely, we found that signals reflecting the spatial category of the sample were stronger and emerged earlier in prefrontal cortex relative to parietal cortex, consistent with a feedback of abstract category signals in the network. This suggests that, within the context of this task, the spatial category signal may originate in prefrontal cortex.

Second, a primary objective of the current experiments was to better understand the neural mechanism by which the brain was able to assign one stimulus to multiple alternative categories according to a variable rule, and we detected a direct neural correlate of this computational flexibility. We found that single neurons exhibited activity that was both category selective and also rule dependent. Neurons of this class were activated when the sample stimulus appeared at a location within the preferred spatial category of each neuron, but only when that category was compatible with the rule in force, in which case the category information carried by neurons was task relevant and provided a basis to select the correct choice (e.g., for neurons preferring either the left or right categories, the LR rule was compatible and AB rule incompatible). This context-dependent activation of neuronal populations coding categories appears to reflect

executive control over a cognitive process, as opposed to a sensorimotor process, in the sense that rule-modulated neurons encoded abstract information (a spatial category) rather than sensorimotor information (such as the position of a stimulus or direction of forthcoming movement). We found that rule modulated category neurons were distributed between prefrontal and parietal cortex, suggesting that the executive control of spatial categorization was a network-level process, involving at a minimum these two cortical areas (and most probably others).

Third, by comparing the strength and timing of rule modulated category signals in the prefrontal-parietal network, we obtained evidence that executive control over spatial categorization in our task may originate within the prefrontal cortex. Specifically, we found that neural signals coding rule-dependent spatial categories (1) were stronger, (2) more deeply modulated by the rule, and (3) emerged earlier in prefrontal cortex relative to parietal cortex. The conclusion that signals reflecting the executive control of categorization were stronger and earlier in prefrontal cortex was strengthened by the fact that these observations were confirmed significant in each of the two monkeys considered individually. However, it is important to note some aspects of executive control as defined by our task were represented by neural signals that emerged in parietal cortex first, notably the signal reflecting the categorization rule, although this signal may have reflected the visual attributes of the boundary cue instructing the rule, rather than an abstract rule per se. Prior work on the neural mechanisms of categorization in the somato-sensory system has shown that neural signals reflecting the category of vibrotactile

stimuli are distributed between the supplementary motor area (SMA) and pre-SMA (Romo et al.,1993, 1997). Prior work in visual categorization has shown that neurons coding object categories based on shape are distributed between prefrontal and infero-temporal cortex (Freedman, Riesenhuber et al. 2001; Freedman, Riesenhuber et al. 2002; Freedman, Riesenhuber et al. 2003; Freedman and Assad 2006; Freedman and Miller 2008; Freedman and Assad 2009) whereas neurons coding categories based on the spatial attributes of visual stimuli are distributed between prefrontal and posterior parietal cortex (Freedman and Assad 2006); (Merchant, Crowe et al. 2011)). Regarding spatial categorization, a recent study compared the timing and strength of category-selective neural signals in parietal and prefrontal cortex of a monkey assigning horizontal bar stimuli to the spatial categories "high" and "low" (Merchant, Crowe et al. 2011). That study and the present one both found that category-selective signals were stronger and arose earlier in prefrontal cortex relative to parietal cortex. Merchant and colleagues found that category signals in parietal cortex emerged only toward the end of the trial near the time of the categorical choice, whereas we observed an earlier recruitment of parietal cortex and a more extensive period of prefrontal-parietal co-activation. A recent study in which monkeys assigned visual stimuli to spatial categories based on their direction of visual motion observed the opposite order of recruitment in the prefrontal-parietal network, with neurons in parietal area LIP exhibiting stronger and earlier category-selective signals relative to their prefrontal counterparts (Swaminathan and Freedman 2012). The reason for the differences in network dynamics observed in these studies is not clear.

One possibility could be that parietal area 7a [the focus of recording in the present study and that of (Merchant, Crowe et al. 2011)]and parietal area LIP [the focus of recording in the study by (Swaminathan and Freedman 2012)] are recruited to represent spatial categories at different times relative to prefrontal cortex. Another possibility is that parietal cortex leads prefrontal cortex in category representation to the degree that visual features and visual categories are tightly coupled, as visually derived information generally flows from parietal to prefrontal cortex. In the present study, spatial position and spatial category were dissociated by requiring that monkeys use two sequentially presented stimuli to compute category, and further by assigning each sample stimulus (and hence each collection of stimulus features) to two alternative categories over trials. Thus, the requirement to compute category using information stored in working memory, or the requirement to do this flexibly in accordance to a changing rule, could shift network dynamics in favor of prefrontal cortex.

Other work on the neural mechanisms of executive control has shown that neurons participating in rule-dependent computations are found both within posterior parietal cortex (Stoet and Snyder 2004) and prefrontal cortex (Wallis, Anderson et al. 2001) (Nakahara, Hayashi et al. 2002). A recent study of particular relevance characterized the neural mechanisms of executive control in prefrontal cortex specifically as it applied to object categorization (Roy, Riesenhuber et al. 2010). Both that prior study and the present one found that, when monkeys had to flexibly assign the same visual

stimulus to competing perceptual categories, the alternative categories were represented by largely distinct populations of prefrontal neurons that were conditionally activated as a function of the category boundary applied. The agreement between these two studies suggests that prefrontal cortex contains a common neural mechanism for flexible categorization that applies to both spatial and object categories. Characterizing differences in the strength and timing of physiological signals that are distributed across interconnected cortical areas provides one approach to disentangling the contributions that each cortical area makes to their shared patterns of electrical activity (Chafee and Goldman-Rakic 1998; Crowe, Chafee et al. 2004; Nieder and Miller 2004). Our present data provide evidence that signals reflecting the interaction between rules and categories, and therefore the executive control of categorization as a cognitive process, are earlier and stronger in prefrontal cortex relative to posterior parietal cortex. These findings may provide a positive image of the cognitive deficit in executive control that has long been recognized as a hallmark of prefrontal cortical damage in humans (Goldman-Rakic 1987; Asaad, Rainer et al. 2000), by providing some of the first single-neuron evidence that prefrontal cortex might be a generator of distributed signals that reflect computational flexibility in cortical networks.

Literature Cited

Andersen, R. A., G. K. Essick, et al. (1985). "Encoding of spatial location by posterior parietal neurons." Science **230**(4724): 456-458.

Asaad, W. F., G. Rainer, et al. (2000). "Task-specific neural activity in the primate prefrontal cortex." J Neurophysiol **84**(1): 451-459.

Averbeck, B. B., D. A. Crowe, et al. (2003). "Neural activity in prefrontal cortex during copying geometrical shapes. II. Decoding shape segments from neural ensembles." Exp Brain Res **150**(2): 142-153.

Bisley, J. W. and M. E. Goldberg (2003). "The role of the parietal cortex in the neural processing of saccadic eye movements." Adv Neurol **93**: 141-157.

Buneo, C. A. and R. A. Andersen (2006). "The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements." Neuropsychologia **44**(13): 2594-2606.

Cavada, C. and P. S. Goldman-Rakic (1989). "Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections." J Comp Neurol **287**(4): 393-421.

Cavada, C. and P. S. Goldman-Rakic (1989). "Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe." J Comp Neurol **287**(4): 422-445.

Cavada, C. and P. S. Goldman-Rakic (1989). "Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe." Journal of Comparative Neurology **287**(4): 422-445.

Chafee, M. V., B. B. Averbeck, et al. (2007). "Representing spatial relationships in posterior parietal cortex: single neurons code object-referenced position." Cereb Cortex **17**(12): 2914-2932.

Chafee, M. V., D. A. Crowe, et al. (2005). "Neural correlates of spatial judgement during object construction in parietal cortex." Cereb Cortex **15**(9): 1393-1413.

Chafee, M. V. and P. S. Goldman-Rakic (1998). "Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task." J Neurophysiol **79**(6): 2919-2940.

Chafee, M. V. and P. S. Goldman-Rakic (1998). "Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task." Journal of Neurophysiology **79**(6): 2919-2940.

Chafee, M. V. and P. S. Goldman-Rakic (2000). "Inactivation of parietal and prefrontal cortex reveals interdependence of neural activity during memory-guided saccades." J Neurophysiol **83**(3): 1550-1566.

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

Crowe, D. A., B. B. Averbeck, et al. (2008). "Neural ensemble decoding reveals a correlate of viewer- to object-centered spatial transformation in monkey parietal cortex." J Neurosci **28**(20): 5218-5228.

Crowe, D. A., B. B. Averbeck, et al. (2010). "Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex." J Neurosci **30**(35): 11640-11653.
Crowe, D. A., B. B. Averbeck, et al. (2005). "Dynamics of parietal neural activity during spatial cognitive processing." Neuron **47**(6): 885-891.
Crowe, D. A., M. V. Chafee, et al. (2004). "Neural activity in primate parietal area 7a related to spatial analysis of visual mazes." Cereb Cortex **14**(1): 23-34.
Freedman, D. J. and J. A. Assad (2006). "Experience-dependent representation of visual categories in parietal cortex." Nature **443**(7107): 85-88.
Freedman, D. J. and J. A. Assad (2009). "Distinct encoding of spatial and nonspatial visual information in parietal cortex." J Neurosci **29**(17): 5671-5680.
Freedman, D. J. and E. K. Miller (2008). "Neural mechanisms of visual categorization: insights from neurophysiology." Neurosci Biobehav Rev **32**(2): 311-329.
Freedman, D. J., M. Riesenhuber, et al. (2001). "Categorical representation of visual stimuli in the primate prefrontal cortex." Science **291**(5502): 312-316.
Freedman, D. J., M. Riesenhuber, et al. (2002). "Visual categorization and the primate prefrontal cortex: neurophysiology and behavior." J Neurophysiol **88**(2): 929-941.
Freedman, D. J., M. Riesenhuber, et al. (2003). "A comparison of primate prefrontal and inferior temporal cortices during visual categorization." J Neurosci **23**(12): 5235-5246.
Funahashi, S., C. J. Bruce, et al. (1989). "Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex." J Neurophysiol **61**(2): 331-349.
Genovesio, A., P. J. Brasted, et al. (2005). "Prefrontal cortex activity related to abstract response strategies." Neuron **47**(2): 307-320.
Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. Handbook of Physiology. The Nervous System. Higher Functions of the Brain. V. B. Mountcastle, F. Plum and S. R. Geiger. Bethesda, MD, Am. Physiol. Soc. **sect. 1, vol. V, chapt. 9:** 373-417.
Goldman-Rakic, P. S. (1995). "Architecture of the prefrontal cortex and the central executive." Ann N Y Acad Sci **769**: 71-83.
Goldman-Rakic, P. S. (1996). "The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive." Philos Trans R Soc Lond B Biol Sci **351**(1346): 1445-1453.
Goldman-Rakic, P. S., J. F. Bates, et al. (1992). "The prefrontal cortex and internally generated motor acts." Curr Opin Neurobiol **2**(6): 830-835.
Greve, K. W., T. R. Stickle, et al. (2005). "Latent structure of the Wisconsin Card Sorting Test: a confirmatory factor analytic study." Arch Clin Neuropsychol **20**(3): 355-364.
Hayashi, T., S. Konishi, et al. (1999). "Short communication: mapping of somatosensory cortices with functional magnetic resonance imaging in anaesthetized macaque monkeys." Eur J Neurosci **11**(12): 4451-4456.
Hoshi, E., K. Shima, et al. (1998). "Task-dependent selectivity of movement-related neuronal activity in the primate prefrontal cortex." J Neurophysiol **80**(6): 3392-3397.

Jo, W., K. J. Freedman, et al. (2012). "Fabrication of tunable silica-mineralized nanotubes using flagella as bio-templates." Nanotechnology **23**(5): 055601.
Johnson, R. A. and D. W. Wichern (1998). Applied Multivariate Statistical Analysis. Saddle River, NJ, Prentice Hall.
Klecka, W. R. (1980). Discriminant Analysis. Newbury Park, CA, Sage Publications, Inc.
Mansouri, F. A., K. Matsumoto, et al. (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**(10): 2745-2756.
Merchant, H., D. A. Crowe, et al. (2011). "Top-down spatial categorization signal from prefrontal to posterior parietal cortex in the primate." Front Syst Neurosci **5**: 69.
Miller, E. K. and J. D. Cohen (2001). "An integrative theory of prefrontal cortex function." Annu Rev Neurosci **24**: 167-202.
Miller, E. K., D. J. Freedman, et al. (2002). "The prefrontal cortex: categories, concepts and cognition." Philos Trans R Soc Lond B Biol Sci **357**(1424): 1123-1136.
Mountcastle, V. B., J. C. Lynch, et al. (1975). "Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space." J Neurophysiol **38**(4): 871-908.
Muhammad, R., J. D. Wallis, et al. (2006). "A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum." J Cogn Neurosci **18**(6): 974-989.
Nakahara, K., T. Hayashi, et al. (2002). "Functional MRI of macaque monkeys performing a cognitive set-shifting task." Science **295**(5559): 1532-1536.
Nieder, A. and E. K. Miller (2004). "A parieto-frontal network for visual numerical information in the monkey." Proc Natl Acad Sci U S A **101**(19): 7457-7462.
Palmer, N. D., C. W. McDonough, et al. (2012). "A genome-wide association search for type 2 diabetes genes in african americans." PLoS One **7**(1): e29202.
Petrides, M. (1985). "Deficits in non-spatial conditional associative learning after periarcuate lesions in the monkey." Behav Brain Res **16**(2-3): 95-101.
Roberts, A. C. and J. D. Wallis (2000). "Inhibitory control and affective processing in the prefrontal cortex: neuropsychological studies in the common marmoset." Cereb Cortex **10**(3): 252-262.
Robinson, D. L., M. E. Goldberg, et al. (1978). "Parietal association cortex in the primate: sensory mechanisms and behavioral modulations." J Neurophysiol **41**(4): 910-932.
Roy, J. E., M. Riesenhuber, et al. (2010). "Prefrontal cortex activity during flexible categorization." J Neurosci **30**(25): 8519-8528.
Snyder, L. H., A. P. Batista, et al. (2000). "Intention-related activity in the posterior parietal cortex: a review." Vision Res **40**(10-12): 1433-1441.
Stoet, G. and L. H. Snyder (2004). "Single neurons in posterior parietal cortex of monkeys encode cognitive set." Neuron **42**(6): 1003-1012.
Stoet, G. and L. H. Snyder (2006). "Effects of the NMDA antagonist ketamine on task-switching performance: evidence for specific impairments of executive control." Neuropsychopharmacology **31**(8): 1675-1681.

Stoet, G. and L. H. Snyder (2009). "Neural correlates of executive control functions in the monkey." Trends Cogn Sci **13**(5): 228-234.
Swaminathan, S. K. and D. J. Freedman (2011). "Distinct roles of parietal and frontal cortices in visual categorization." 2011 Neuroscience Meeting Planner. **324.05**.
Swaminathan, S. K. and D. J. Freedman (2012). "Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex." Nat Neurosci **15**(2): 315-320.
Tanji, J. and E. Hoshi (2008). "Role of the lateral prefrontal cortex in executive behavioral control." Physiol Rev **88**(1): 37-57.
Tsujimoto, S., A. Genovesio, et al. (2008). "Transient Neuronal Correlations Underlying Goal Selection and Maintenance in Prefrontal Cortex." Cereb Cortex.
Wallis, J. D., K. C. Anderson, et al. (2001). "Single neurons in prefrontal cortex encode abstract rules." Nature **411**(6840): 953-956.