Spiking and oscillatory correlates of visual short-term memory for multiple items

by

Simon Kornblith

B.S., California Institute of Technology (2010)

Submitted to the Department of Brain and Cognitive Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

Author. Signature redacted .....
Department of Brain and Cognitive Sciences
May 5, 2017

Signature redacted

Certified by .....
Earl K. Miller
Picower Professor of Neuroscience
Thesis Supervisor

Signature redacted

Accepted by .....
Matthew Wilson
Sherman Fairchild Professor of Neuroscience and Picower Scholar
Director of Graduate Education for Brain and Cognitive Sciences

Spiking and oscillatory correlates of visual short-term memory for multiple items

by

Simon Kornblith

Submitted to the Department of Brain and Cognitive Sciences
on May 5, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The richness of visual experience far exceeds our ability to remember what we have seen. However, it is unclear what neural mechanisms give rise to these limits to visual short-term memory capacity. Here, we measured neural activity in a change localization task, in which monkeys viewed two displays of multiple colored squares separated by a brief delay, and made a saccade to the square that changed color between displays. In chapter 2, we examine local field potentials in the lateral intraparietal area (LIP), frontal eye field, and lateral prefrontal cortex (PFC). At stimulus encoding, lower frequency oscillations decreased in power in proportion to the total number of stimuli presented, while higher frequency oscillations increased in power in proportion to the number of stimuli contralateral to the recording site. During the delay, lower frequency power instead increased with the number of contralateral stimuli, while higher frequency power was not modulated. We interpret these findings in terms of roles for low- and high-frequency oscillations in changing and maintaining cognitive state. In chapter 3, we compare spiking activity between LIP, PFC, and inferotemporal cortex (IT). Although the task required that the animal remember stimulus colors, activity in LIP and PFC primarily reflected the stimulus positions, while activity in IT primarily reflected color. In PFC, color information increased with the number of stimuli presented, while in IT, color information remained constant or decreased. Thus, IT was more strongly capacity-limited than PFC. Color selectivity during the delay was weak in all regions. However, in IT, activity at test stimulus presentation reflected the difference in square colors between the sample and test displays, while in PFC, activity primarily reflected the location of the changed square. Selectivity to these attributes was stronger on correct trials than incorrect trials. Our findings suggest a possible role for passive processes in IT in visual short-term memory.

Thesis Supervisor: Earl K. Miller
Title: Picower Professor of Neuroscience

Acknowledgments

I thank my advisor, Earl Miller, for supporting me throughout my graduate career. Our discussions have been extremely useful in helping me to understand the weird world of monkeys, cognition, and prefrontal cortex. You have taught me how to design compelling experiments and build a compelling story. I appreciate your ability to challenge my views while still encouraging me to think independently.

I am also greatly indebted to my undergraduate mentors Florian Mormann and Doris Tsao. Florian first introduced me to the world of neuroscience research, taught me how to analyze electrophysiological data, and impressed upon me the value of proper statistical methods. Doris was willing to take a leap of faith and give me, then an undergraduate, a monkey electrophysiology project of my own. She taught me how to think about vision and the brain, how to work with monkeys, and how to navigate academia. Our conversations continue to guide me both personally and professionally. I am immensely fortunate that, at each stage of my early scientific career, my mentors treated me as their scientific equal, and encouraged me to learn all that I could about the things that interested me.

This research would not have been possible without the intellectual and technical contributions and personal support of my colleagues in the Miller Lab, especially Tim Buschman, Scott Brincat, Jacob Donoghue, Roman Loonis, Mikael Lundqvist, Jonas Rose, and Jefferson Roy. You were always willing to help me in whatever way you could, and encouraged me to persist despite the many setbacks I encountered. I would also like to express my gratitude to the managers of the Miller Lab, Alexa D'Ambra, Brenna Gray, and Diana Ouellette, for their efforts to ensure that I had all the resources I needed to perform my experiments.

Finally, I thank my parents, Carol and Gary, my brothers, Max and Russell, and my partner Aliza, for their love and support. You believed in me throughout this process, and I owe everything to you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cognitive Capacity, Short-Term Memory, and Magic Numbers

The idea of an intrinsic cognitive capacity limit was first formulated by G. A. Miller (1956). Miller observed that, in studies of absolute judgments of pitch, loudness, saltiness, and position on a line, the mutual information between the stimulus and judgment was remarkably consistent, averaging 2.6 bits, or 6.5 categories. Moving on to studies of short-term memory span for letters, digits, and words, he observed that the number of items that subjects could recall was similarly limited to 7 +/- 2, regardless of the complexity of the individual items. While further studies have downgraded the standard estimate of short-term memory capacity to 3 to 5 items (Henderson, 1972; Luck & Vogel, 1997; Cowan, 2001), both the concept of "capacity," whether of cognition in general and of short-term memory in particular, and the idea of a "magic number" that indexes the limit to this resource continue to guide many experiments.

In change detection tasks, which require subjects to detect whether a display of items has changed across a brief delay, studies have reported no effect of the type of stimuli (Luck & Vogel, 1997) or the duration of the sample display or the delay (Bays, Gorgoraptis, Wee,

Marshall, & Husain, 2011; Tsubomi, Fukuda, Watanabe, & Vogel, 2013) upon subsequent behavioral performance. Although some authors have interpreted these findings as evidence for a single, fixed capacity limit for maintaining visual information, recent studies have challenged this assertion. Visual search speed and short-term memory performance appear to be higher for simple objects compared to complex objects, suggesting that the amount of information to be maintained, rather than the total number of items, controls the precision of visual short-term memory representation (Wheeler & Treisman, 2002; Alvarez & Cavanagh, 2004). However, advocates of fixed capacity have proposed that the apparent reduction in capacity for complex objects result from increased sample-test display similarity, rather than object complexity *per se* (Awh, Barton, & Vogel, 2007). Moreover, Brady, Störmer, and Alvarez (2016) recently showed that capacity is not independent of the temporal parameters of the task. They found that observed insensitivity to encoding time holds only for artificial stimuli, such as colored squares. When subjects maintained images of objects, performance increased with exposure duration. EEG activity during the delay was greater for trials when the exposure duration was longer, suggesting that the behavioral advantage reflects active maintenance processes, rather than use of long-term memory. Thus, although the capacity of visual short-term memory appears to be consistent under a wide variety of conditions, a single "magic number" cannot explain performance in all tasks.

1.2 Models of Multiple-Item Short-Term Memory

The finding that performance in change detection tasks achieves ceiling levels at loads of 3 to 4 items, but drops as load is further increased, led Luck and Vogel (1997) to postulate that our visual short-term memory "store" has 3 to 4 discrete "slots" in which items may be held. According to this view, when the number of items to be remembered exceeds the number of available slots, information about some items must be discarded. A competing view proposes that visual short-term memory is a shared resource that is distributed among all items in the stimulus display, and performance decreases when more than 3 to 4 items

must be remembered because an insufficient amount of this resource can be devoted to each individual item.

Studies of visual short-term memory precision using continuous report paradigms have shown that neither of these models are entirely correct. In these studies, as in change detection studies, subjects see an array of multiple colored squares, which is followed by a delay period. However, instead of signaling whether a second display contains a change, subjects report the color of a specific item in the initial display by clicking on a color wheel. A naïve slot model predicts that the distribution of the subjects' choices around the correct response should be the same for all stimulus displays below capacity, but when above capacity, the distribution should be a mixture of the below capacity distribution and a uniform distribution that corresponds to guessing of items not encoded. A naïve resource model predicts that the response distribution should be "stretched" as the number of items increases regardless of whether the initial display was above or below capacity. In reality, the response distribution is stretched as the number of items increases up to capacity, but more closely resembles a mixture distribution when the number of items exceeds capacity (Zhang & Luck, 2008; Gorgoraptis, Catalao, Bays, & Husain, 2011). Zhang and Luck (2008) proposed that the discrepancy with the naïve slot model could be explained by a "slots + averaging" or "discrete representation" model, in which the number of slots remains fixed, as in the ordinary slot model, but multiple slots may be devoted to the same object. However, Bays, Catalao, and Husain (2009) showed that incorrect responses were not random, but instead clustered around the colors of non-target squares, consistent with a resource model in which both color and position information must compete for representation.

Additional studies have shown that subjects' behavior can be readily explained by a "variable precision" resource model, in which the division of resources among items in the stimulus display is not equal, but instead varies from trial to trial (Fougnie, Suchow, & Alvarez, 2012; van den Berg, Shin, Chou, George, & Ma, 2012; Ma, Husain, & Bays, 2014). The variable precision model explains two additional qualities of observed experimental results. First, the uniform-normal mixture proposed by Zhang and Luck (2008) is not a particularly

good fit to behavior in studies of visual short-term memory precision. The variable precision model implies that the distribution of responses should be more closely approximated by an infinite mixture of normal distributions of different variances, which is indeed the case (Fougnie et al., 2012; van den Berg et al., 2012). Second, subjects have some metacognitive information regarding which items were encoded better than others, and subjects can recall the item they believe they encoded best better than other items (Fougnie et al., 2012).

1.3 Neural Mechanisms of Visual Short-Term Memory

Most single neuron electrophysiology studies focus on persistent, stimulus-selective spiking during the delay period as the neural substrate of visual short-term memory. Robust stimulus-selective persistent delay period activity for visual stimuli has been reported in the inferotemporal cortex (IT) (Fuster & Jervey, 1982; Miyashita & Chang, 1988; E. K. Miller, Li, & Desimone, 1993) and prefrontal cortex (PFC) (Fuster & Alexander, 1971; Niki & Watanabe, 1976; F. A. W. Wilson, Scalaidhe, & Goldman-Rakic, 1993; E. K. Miller, Erickson, & Desimone, 1996). Additionally, several studies have reported that lesions to or reversible inactivation of these regions produce deficits in visual short-term memory performance (Passingham, 1975; Fuster, Bauer, & Jervey, 1985; Horel, Pytko-Joiner, Voytko, & Salsbury, 1987; Kowalska, Bachevalier, & Mishkin, 1991; Gaffan & Murray, 1992). However, these lesion studies are somewhat difficult to interpret. Impairments are often relatively small. Deficits arising from IT lesions could reflect impaired ability to process visual stimuli, whereas deficits arising from PFC lesions could reflect impaired executive control processes. One study reported that sufficiently large lesions to ventrolateral prefrontal cortex impaired both simultaneous and delayed color discrimination, but all deficits disappeared after animals were retrained to preform the simultaneous matching task (Rushworth, Nixon, Eacott, & Passingham, 1997).

Stimulus-selective delay period activity is an intuitive way in which information about previously presented stimuli could be maintained in neurons, but recent studies suggest that

it may not be necessary or sufficient. Rigotti et al. (2013) showed that, although sample stimulus identity could be readily decoded from PFC neurons during the delay period even on incorrect trials with nearly perfect accuracy, the number of implementable binary classifications between different stimulus conditions was substantially lower on incorrect trials. Thus, beyond the mere presence or absence of persistent activity, the dimensionality of such activity appears to be important for behavior. Other recent work has emphasized the importance of short-term synaptic plasticity in maintaining information during the delay (Stokes, 2015). Although modeling studies have confirmed the adequacy of short-term plasticity to support short-term memory (Mongillo, Barak, & Tsodyks, 2008; Lundqvist, Herman, & Lansner, 2011), it is difficult to observe directly. Thus, at present, evidence for a truly "activity-silent" representation of the contents of short-term memory is somewhat lacking. The original motivation for this hypothesis came from a study in which animals performed a secondary task while maintaining a stimulus in short-term memory. The magnitude of stimulus-selective persistent activity dropped during this second task, and then rebounded when it was complete (Watanabe & Funahashi, 2014). Although this study underscores the dynamic nature of prefrontal representations, persistent activity was never totally absent. Rose et al. (2016) recently reported that they could "reactivate" a latent representation of a memory item that was not presently task-relevant using transcranial magnetic stimulation (TMS). Before TMS, the unattended memory identity could not be decoded from EEG, but TMS produced a small but significant increase in decoding accuracy above chance. However, there are alternative potential explanations for this effect. If persistent activity is present but weak, it might not be decodable simply because the signal is insufficiently strong. TMS might act by boosting this pre-existing persistent activity or suppressing intrinsic neuronal variability, rather than by revealing otherwise silent representations in synaptic weights.

Recent reports indicate that early visual areas may also play a role in short-term memory maintenance. fMRI studies have shown that it is possible to decode remembered stimuli from areas as early as V1 (Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009). However, electrophysiological evidence for delay period activity in early visual areas is mixed.

Although some studies have found robust delay period activity in V1 (Supèr, Spekreijse, & Lamme, 2001; van Kerkoerle, Self, & Roelfsema, 2017) and V4 (Hayden & Gallant, 2013), other studies have reported only weak modulation in a small proportion of neurons (Chelazzi, Miller, Duncan, & Desimone, 2001; Lee, Simpson, Logothetis, & Rainer, 2005). Discrepancies between these studies could reflect differences in task designs and training procedures, but it is also possible that observed fMRI results reflect factors other than delay period activity, as fMRI signals can diverge from spiking in some circumstances (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Maier et al., 2008; Sirotin & Das, 2009). In particular, Lee et al. (2005) have shown that the contents of visual short-term memory modulate the strength of ongoing theta oscillations in V4 as well as the degree of phase locking to these theta oscillations.

1.4 Neural Mechanisms of Capacity Limitations

A large body of research implicates the posterior parietal cortex (PPC) in capacity limitations. fMRI studies have shown that the BOLD response in the human intraparietal sulcus (IPS) increases with stimulus load up to an individuals' behavioral capacity, then saturates once capacity is reached (Todd & Marois, 2004, 2005; Xu & Chun, 2005; Robitaille et al., 2010). Further research has differentiated the BOLD response in the superior IPS, which appears to track the total number of objects regardless of complexity, from the inferior IPS, where the BOLD response is more strongly modulated by object complexity (Xu & Chun, 2005; Xu, 2007, 2008). In addition, an event-related potential, termed "contralateral delay activity" (CDA), can be measured by recording occipitoparietal EEG electrodes while subjects perform a visual short-term memory task in which they are cued to remember items from the contralateral side. Similar to BOLD responses in the IPS, CDA increases with stimulus load contralateral to the recording site and saturates at individuals' behavioral capacity (Vogel & Machizawa, 2004; McCollough, Machizawa, & Vogel, 2007). However, the PPC is unlikely to be the only region implicated in capacity limitation. Several fMRI studies have

also reported modulation of the BOLD response with load in other regions, including PFC, occipitotemporal cortex, and the anterior cingulate (Linden et al., 2003; Todd & Marois, 2004). Furthermore, human subjects with lesions to either the basal ganglia or PFC show impaired performance in multiple-item visual short-term memory tasks and no load-related modulation in CDA (Voytek & Knight, 2010).

In the last five years, several studies have investigated multiple item visual short-term memory and visual short-term memory capacity by invasive methods in non-human primates using similar tasks to those previous reports in humans. Buschman, Siegel, Roy, and Miller (2011) showed that the amount of information about each encoded item decreased in the lateral intraparietal area (LIP), frontal eye fields (FEF), and PFC as the number of items presented in the contralateral hemifield increased. However, this study used a simple linear model to measure encoded information about each stimulus, whereas neurons could potentially respond non-linearly to different stimulus combinations. Lara and Wallis (2014) recorded from PFC in a one- or two-item change detection task with equiluminant colored squares. They reported that, even when only a single colored square was presented, scarcely more PFC neurons represented information about the color to be remembered than would be expected by chance.

While the studies above have revealed regions putatively implicated in capacity limits and shown that these limits manifest themselves in neural responses, little is known about the source of the capacity limit itself. One proposal from Lisman and Idiart (1995) suggests that neurons representing specific stimuli are sequentially activated every gamma cycle, with the whole sequence of remembered stimuli repeated every theta cycle. Although there is little evidence for this specific model, the idea that oscillatory phase could separate multiple items held in short-term memory remains influential. Several studies of oscillatory activity and short-term memory have shown parametric modulation by load (Jensen & Tesche, 2002; Howard et al., 2003; Meltzer et al., 2008; van Vugt, Schulze-Bonhage, Litt, Brandt, & Kahana, 2010; J. M. Palva, Monto, Kulashekhar, & Palva, 2010; S. Palva, Kulashekhar, Hämäläinen, & Palva, 2011; Roux, Wibral, Mohr, Singer, & Uhlhaas, 2012). However, sim-

ple modulation of oscillatory activity provides only weak evidence for phase coding. Other studies provide more explicit evidence. Siegel, Warden, and Miller (2009) recorded from monkey PFC during a two item visual short-term memory task with sequentially presented stimuli. The authors showed that spikes encoding information about the first item were concentrated in earlier phases of the ongoing gamma oscillation, while spikes encoding information about the second item were concentrated in later phases. Axmacher et al. (2010) report that cross-frequency coupling between theta phase and gamma amplitude in human hippocampus increases with stimulus load. While these studies alone each provide evidence for phase coding, it is somewhat discomforting that they report very different effects: Siegel et al. (2009) imply sequence replay within gamma cycles, whereas Axmacher et al. (2010) imply sequence replay within theta cycles. While it is possible that these discrepancies arise from true differences in the brain areas investigated, further studies are necessary to confirm that phase coding plays a role in multiple item visual short-term memory.

Other models have suggested that capacity limits arise not from limits in coding within oscillatory cycles, but from normalization. Normalization refers to the finding that the firing rates of neurons are intrinsically limited by the firing of neighboring neurons, such that the summed activity within a group remains constant (Carandini & Heeger, 2012). A substantial literature exists regarding normalization in neural circuits from the *Drosophila* olfactory bulb to low- and mid-level visual areas (Reynolds & Heeger, 2009; Olsen, Bhandawat, & Wilson, 2010; N. R. Wilson, Runyan, Wang, & Sur, 2012). Both computational (Wei, Wang, & Wang, 2012) and statistical (Bays, 2014) models have recently been proposed that reproduce human behavioral capacity results by including normalization in population coding of remembered stimuli. Electrophysiological (Buschman et al., 2011; Balan, Oristaglio, Schneider, & Gottlieb, 2008; Mirpour & Bisley, 2012) and fMRI (Sprague, Ester, & Serences, 2014) studies have shown that, when multiple stimuli are presented simultaneously, neural responses to any given stimulus in prefrontal, parietal, and early visual cortices are weakened. Further studies are necessary to determine whether measured normalization effects accurately reflect behavioral capacity limitations.

Chapter 2

Stimulus Load and Oscillatory Activity in Higher Cortex

The contents of this chapter are reproduced from Simon Kornblith, Timothy J. Buschman, and Earl K. Miller, Stimulus load and oscillatory activity in higher cortex, *Cerebral Cortex*, 2016, 26(9), 3772–3784, by permission of Oxford University Press.

2.1 Abstract

Exploring and exploiting a rich visual environment requires perceiving, attending, and remembering multiple objects simultaneously. Recent studies have suggested that this mental "juggling" of multiple objects may depend on oscillatory neural dynamics. We recorded local field potentials from the lateral intraparietal area, frontal eye fields, and lateral prefrontal cortex while monkeys maintained variable numbers of visual stimuli in working memory. Behavior suggested independent processing of stimuli in each hemifield. During stimulus presentation, higher frequency power (50-100 Hz) increased with the number of stimuli (load) in the contralateral hemifield, while lower frequency power (8-50 Hz) decreased with the total number of stimuli in both hemifields. During the memory delay, low frequency power increased with contralateral load. Load effects on higher frequencies during stimulus en-

coding and lower frequencies during the memory delay were stronger when neural activity also signaled the location of the stimuli. Like power, higher frequency synchrony increased with load, but beta synchrony (16-30 Hz) showed the opposite effect, increasing when power decreased (stimulus presentation) and decreasing when power increased (memory delay). Our results suggest roles for lower frequency oscillations in top-down processing and higher frequency oscillations in bottom-up processing.

2.2 Introduction

While a wealth of studies have investigated how humans and animals remember single objects, less is known about how we remember multiple objects simultaneously. It is well-known that human capacity for multiple items is severely limited: The average adult human can only hold 3-4 objects in mind (working memory) simultaneously (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001). fMRI signals from a variety of brain areas are modulated by the number of remembered stimuli (stimulus load) (Linden et al., 2003; Todd & Marois, 2004). Similarly, EEG event-related potentials scale with stimulus load, saturate when behavioral capacity is exceeded, and reflect individual differences in working memory capacity (Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005; McCollough et al., 2007; Ikkai, McCollough, & Vogel, 2010; Luria & Vogel, 2011). However, multiple-item working memory processes are much less studied in animals, where we can assess neural activity with greater temporal and spatial precision.

We previously reported correlates of multiple-item working memory in individual neurons in frontal and parietal cortex (Buschman et al., 2011). Monkeys performed a human test of capacity (Figure 2-1). Two arrays of 2-5 colored squares were separated by a memory delay. The color of a random square was changed. Monkeys were trained to saccade to this change. Multiple electrodes were implanted in lateral prefrontal cortex (lPFC), frontal eye fields (FEF), and lateral intraparietal area (LIP). These regions are critical for short-term memory (Passingham, 1975; Kowalska et al., 1991; Sawaguchi & Goldman-Rakic, 1991;

Li, Mazzoni, & Andersen, 1999) and human studies implicate them in capacity limitations (Linden et al., 2003; Todd & Marois, 2004, 2005; Vogel & Machizawa, 2004; J. M. Palva et al., 2010; Voytek & Knight, 2010). On the neuron level, we found capacity limitations were bottom-up (appearing in parietal before frontal cortex), neural information about the target stimulus decreased with stimulus load even when these stimuli was correctly remembered, and neural information was present but reduced in trials in which the animal failed to select the correct target (Buschman et al., 2011).

Here, we use these data to examine effects of stimulus load on oscillations of local field potentials (LFPs). There is increasing evidence that oscillations play a role in cognition. Different oscillatory frequencies may mediate feedforward versus feedback processing (Engel, Fries, & Singer, 2001; Buschman & Miller, 2007; Engel & Fries, 2010; Arnal & Giraud, 2012; Bastos et al., 2015) and dynamically link neurons into ensembles (Gray, König, Engel, & Singer, 1989; Buschman, Denovellis, Diogo, Bullock, & Miller, 2012; Salazar, Dotson, Bressler, & Gray, 2012). Capacity limits have been hypothesized to arise from coding of different stimuli at different oscillatory phases (Lisman & Idiart, 1995; Siegel et al., 2009; Lundqvist et al., 2011). The few human EEG studies investigating oscillatory activity showed that power and synchrony increases with stimulus load across a range of frequencies (Jensen & Tesche, 2002; Howard et al., 2003; Meltzer et al., 2008; J. M. Palva et al., 2010; S. Palva et al., 2011). But the link between oscillations and multiple item working memory is not well-studied and thus far from understood.

2.3 Methods

2.3.1 Behavioral Task

One adult male rhesus macaque (Macaca mulatta) and one adult male cynomolgus macaque (Macaca fascicularis) were trained to perform a change localization task. All procedures followed the guidelines of the Massachusetts Institute of Technology Committee on Animal Care and the National Institutes of Health. Animals fixated for 500 ms to initiate a trial.

After this fixation period, an array of 2-5 colored squares (1-3 per hemifield) appeared for 800 ms. The stimuli then disappeared. After a 800-1000 ms memory delay period, the array reappeared with a change to the color of the random square. The animal received a juice reward for making a direct saccade to the changed square. From the start of the trial until the presentation of the second array of colored squares, the animal was required to fixate within 1.75 degrees of a central fixation point. Task timing was similar to a previous human study (Luck & Vogel, 1997). In humans, retinal persistence of stimuli presented for 500 ms is approximately 100 ms after stimuli offset (Coltheart, 1980), so it is unlikely that the task could be accomplished on the basis of retinal persistence alone. When one of the authors (SK) performed the task, there was no indication of retinal persistence that substantially outlasted stimulus presentation.

The location of the target was randomized for each trial. However, in order to permit inspection of neural encoding of visual information during the delay period, square locations were chosen from 6 positions (3 per hemifield) in any single session, and only 2 colors could be present at any single position. Additionally, given our behavioral evidence for the independence of working memory representations in each hemifield, in each trial, we manipulated the number of stimuli in each hemifield rather than the total number of stimuli in the display, while constraining the total number of stimuli to between two and five. In order to maintain behavioral performance and to acquire a sufficient number of trials with low total load, trials with fewer stimuli in a given hemifield were presented at higher probability than trials with greater numbers of stimuli. This trial selection procedure was uninformative about the location of the target and did not provide any additional information that could assist in the performance of any given trial.

New stimulus locations and colors were randomly selected before each recording session. Stimulus locations were selected to be within 75 degrees of polar angle from the horizontal meridian and between 4-6 degrees of visual angle from the central fixation point, and colors were manually inspected to ensure sufficient discriminability. All twelve possible colors in a given session were unique. The minimum difference between the sample and test color of

the changed square was CIEDE2000 $\Delta E_{00} = 23.4$ (mean $\Delta E_{00} = 55.6$; max $\Delta E_{00} = 113.8$). Thus, color changes were very perceptible to a human observer (Yang, Ming, & Yu, 2012). An infrared video eye tracking system (ISCAN) recorded eye positions at 240 Hz. A computer running the MonkeyLogic software (Asaad & Eskandar, 2008) controlled the stimulus display and behavioral reward.

2.3.2 Electrophysiological Recording

Epoxy-coated dura-piercing tungsten electrodes (FHC) were lowered into each region using a custom-built grid and microdrive assembly that lowered electrodes in pairs using a single screw. Recordings were performed using a Plexon Multichannel Acquisition Processor. All signals were referenced to ground. Local field potentials were filtered with hardware filters between 3.3 Hz and 88 Hz prior to amplification and sampled at 1000 Hz. 2-pole notch filters at 60 Hz (line noise frequency), 85 Hz (monitor refresh rate), and 120 Hz (line noise harmonic) were applied prior to data analysis.

We restricted our analyses to electrodes from which we recorded at least one single unit. Only parietal electrodes whose units showed a spatially selective response in a delayed saccade task ($p < 0.05$, permutation test of ω^2) were included in further analyses. Electrodes in FEF and lPFC were differentiated using microstimulation. Out of 1125 recorded electrodes, 546 electrodes fit our criteria (2.1).

Region	Total	Mean	IQR	Min	Max
LIP	142	5.1	3.75–7.25	1	9
FEF	155	5.5	4.75-6	1	12
PFC	249	8.9	7-10.25	5	15
All	546	19.5	18-21.25	13	28

Table 2.1: Number of electrodes analyzed for main analyses. "Total" denotes the total number of penetrations in each region. "Mean," "IQR," "Min," and "Max" denote the mean, inter-quartile range, minimum, and maximum numbers of electrodes in individual sessions.

2.3.3 Estimation of Behavioral Capacity

Our procedure for estimating behavioral capacity by mutual information is described in detail in Buschman et al. (2011). To determine mutual information for the entire display, for each load, we computed the conditional mutual information between the animal's choice and the target given the stimulus display:

$$I(\text{choice; target}|\text{display}) =$$
$$\sum_{x \in \text{displays}} \sum_{y=1}^{6} \sum_{x=1}^{6} P_{\text{choice,target,display}}(x, y, z) \log \left(\frac{P_{\text{display}}(z) P_{\text{choice,target,display}}(x, y, z)}{P_{\text{choice,display}}(x, z) P_{\text{target,display}}(y, z)} \right)$$

where x represents the animals' chosen square, y represents the correct target, and indicates each possible display for a given load condition. The resulting values are shown in Figure 2-1C. To determine mutual information for each hemifield, for each stimulus display, we computed the mutual information between the choice and target for each individual display:

$$I(\text{choice; target}) = \sum_{y=1}^{n} \sum_{x=1}^{n} P_{\text{choice,target}}(x, y) \log \left(\frac{P_{\text{choice,target}}(x, y)}{P_{\text{choice}}(x) p_{\text{target}}(y)} \right)$$

Since the behavioral analysis in Figure 2-1B showed that the two hemifields were independent, we dissociated the amount of mutual information in each hemifield using linear regression, fitting the mutual information for the $\sum_{i=2}^{6} \binom{n}{i} = 56$ possible displays (15 2-stimulus displays, 20 3-stimulus displays, 15 4-stimulus displays, and 6 5-stimulus displays) against a design matrix of 16 indicators representing each of the $\sum_{i=0}^{3} \binom{6}{i} = 8$ possible left and right sides. The resulting coefficients averaged by load are shown in Figure 2-1D. Confidence intervals were computed by bootstrapping individual sessions.

Because we asked subjects to localize a changed object, rather than detect whether an object has changed, we cannot measure capacity in this task using Pashler's and Cowan's K, which are the most commonly reported measures in human studies. Previous human localization using a task similar to ours have reported either K = (proportion correct $\times$ set size) $- 1$ (Gold et al., 2006) or K = proportion correct $\times$ set size (Johnson et al., 2013).

When applied to our data, the latter measure yields capacity estimates of $K = 1.6$ (SEM = 0.02) at total load 2, $K = 2.1$ (SEM = 0.03) at load 3, $K = 2.8$ (SEM = 0.04) at load 4, and $K = 3.1$ (SEM = 0.06) at load 5. $K = 2.8$ at load 4 falls within the reported range for healthy human subjects (Johnson et al., 2013).

2.3.4 Time-Frequency Analysis

All analyses were performed using the Julia programming language (http://www.julialang.org/). Spectrograms and coherograms in Figures 2-2, 2-3, 2-6, and 2-S4 were computed by continuous wavelet transform. The continuous wavelet transform was computed using complex Morlet wavelets by multiplication of the Fourier transform of the data with the analytically computed Fourier transform of the wavelet (Torrence & Compo, 1998). Our 22 daughter wavelets had center frequencies spanning from 2.8 Hz to 107.6 Hz and a multiplicative step of $\sqrt[4]{2}$ between frequencies with a wavenumber of 5. While the highest frequencies we investigated exceed the cutoff of the hardware filter, the cutoff is not hard (58% power reduction at 100 Hz), and our analyses are sensitive only to relative changes in power. We computed wavelet transforms for all samples recorded from each electrode and session, and then cut the data according to task epochs and preserved the wavelet coefficients corresponding to every tenth time point for further analysis. We linearly interpolated time points at which the amplifier saturated and removed trials where the amplifier saturated for 5 or more successive samples.

Before computing power and phase statistics, we removed the evoked potentials by analyzing the sum of the squares of the real and imaginary components of the wavelet coefficients after fitting and removing signals that were phase-locked across trials. We fit the real and imaginary components of the wavelet coefficients of each frequency and time point using a design matrix comprising indicators for each contralateral and ipsilateral load condition and performed subsequent analysis on the least squares residuals. Since an intercept column would be linearly dependent with the columns in this design matrix, subtracting the fit components also removes evoked potentials unaffected by load. Because the real and imaginary

components of the wavelet coefficients are sums of the original samples with fixed weights, this procedure is equivalent to regressing the load conditions from the original samples before computing the wavelet transform. This procedure leaves only induced potentials, which are by definition of random phase across trials.

Bar and line graphs in Figures 2-4, 2-5, 2-7, 2-S1, 2-S3, and 2-S4 were computed using multitaper time-frequency transforms. The time-bandwidth product NW was tailored to the specific epoch and frequency band as:

$$\frac{(t_{\text{end}} - t_{\text{start}})(f_{\text{end}} - f_{\text{start}})}{2}$$

where t_{start} and t_{end} were the first and last time points of the signal in seconds, f_{start} and f_{end} were the first and last frequencies in the band. Only the $\lceil 2NW - 1 \rceil$ maximally concentrated discrete prolate spheroidal sequences were used as tapers. As for the wavelet transforms, the evoked potential was removed by linear regression of each ipsilateral and contralateral load condition.

2.3.5 Oscillatory Power Analysis

Power was modeled with a generalized linear model with a gamma likelihood function and a logarithmic link function. We chose to use a gamma likelihood because it yields the maximum likelihood estimator of the effects under the assumption that the wavelet coefficients follow a circularly symmetric complex normal distribution. Under this assumption, the wavelet power is distributed as $|W|^2 \sim \frac{1}{2} P_k \chi^2_2$, where W is the estimated wavelet coefficient and P_k is the power for the given stimulus conditions (Torrence & Compo, 1998). Equivalently, $|W|^2 \sim \text{Gamma}(1, P_k)$, where the gamma distribution is parameterized by its shape and scale. Note that this implies that the standard deviation of the power increases linearly with its mean. A similar relationship applies to the multitaper coefficients. Because each taper provides an orthogonal estimate of power, the average power estimated with n tapers is distributed as $\frac{1}{n} \sum_{i=1}^{n} |M_i|^2 \sim \text{Gamma}\left(n, \frac{P_k}{n}\right)$, where M_i is the coefficient for the ith taper.

We chose the logarithmic link to give a convenient interpretation of the power, although fitting models with the identity link yielded similar results and comparable values for the deviance. All plots show mean parameter values transformed to percent signal change as $\mathrm{PSC}(\beta) = 100(e^{\beta} - 1)$.

We compared the gamma distribution model of power changes with load to an ordinary linear (normal) model, as well as a linear model applied to the log-transformed power. The likelihood of the gamma model was greater than that of the normal model for 98% of recorded electrodes, frequencies, and time points, and greater than that of the log-transformed normal model for 78% of electrodes, frequencies, and time points, indicating that it is a better fit to the data.

Because this model reasonably approximates the underlying distribution of the data, significance and confidence intervals for individual parameters can be computed without resampling. However, because signals in different electrodes recorded simultaneously could be correlated, we computed confidence intervals and p-values by bootstrapping trials recorded in each session and fitting a normal distribution to the mean of these bootstraps across all sessions weighted by $n_{\text{electrodes}} \times n_{\text{trials}}$. The normality assumption is justified by the central limit theorem. We used 100 bootstraps for the time-frequency analysis in Figures 2-2 and 2-3, and 1000 bootstraps for the bar graphs in Figure 2-4 and p-values reported in the text.

2.3.6 Permutation Test for Monotonic Relationship Between Power and Load

To determine whether power and synchrony changes were strictly monotonic, we applied a permutation test. We first determined whether the estimated power for each load condition was strictly monotonic, i.e., $\mathrm{sign}(P_2 - P_1) = \mathrm{sign}(P_3 - P_2)$, where P_n is the average power in a given band and epoch across all recorded electrodes. If this condition was satisfied, we then computed the test statistic $\hat{T} = T(P_1, P_2, P_3) = \min(|P_2 - P_2|, |P_3 - P_2|)$, the minimum difference in power between any two load conditions. If this condition was not satisfied, we set $\hat{T} = 0$. To determine the distribution of $\hat{T}$ under the null hypothesis, we permuted load

condition labels 25,000 times and computed T for power estimates based on the permuted labels. We then calculated the p-value as the proportion of permutations for which $T > \hat{T}$. The significance threshold in Figures 2-5 and 2-S4 was fixed at $p < 0.05$ after Holm multiple testing correction.

2.3.7 Assessing Position Information in Power

We used the gamma generalized linear model described above to assess information about the locations of presented stimuli in measured oscillatory power. Because load information is nested in position information (i.e., stimulus load can be described as the sum of the indicators representing individual positions), we determined whether significant position information was present by model comparison. The full model incorporated indicators for all 6 possible stimulus positions and the contralateral and ipsilateral load zero conditions, as well as an intercept. We assessed whether position information was present for a given hemifield by comparing whether the full model explained more deviance than a null model where the position indicators in that hemifield were replaced with a load regressor representing their sum. Since these two models are nested:

$$\frac{\hat{v}}{k_{\text{full}} - k_{\text{null}}} \left(D(y; \mu_{\text{null}}) - D(y; \mu_{\text{full}})\right) \sim F_{k_{\text{full}}-k_{\text{null}}, n-k_{\text{full}}}$$

(Fox, 2008). $D(y;\mu) = -2\sum_{i=1}^{n}\left(\log\left(\frac{y_i}{\mu_i}\right) - \frac{y_i-\mu_i}{\mu_i}\right)$, the gamma model deviance, k_{full} is the number of predictors in the full model, k_{null} is the number of predictors in the null model, and $\hat{v}$ is the shape parameter of the gamma distribution estimated by the moment estimator $\hat{v}^{-1} = \frac{\sum_{i=1}^{n}(y_i-\hat{\mu}_i)^2/\hat{\mu}_i^2}{n-k}$, where y is the vector of observed values, $\hat{\mu}$ is the vector of fitted values, and k is the number of predictors in the model. (Although v is known *a priori* under the assumptions above, and thus a χ^2 test could be used in place of the F-test, we nonetheless estimate v from the data to account for possible over- or under-dispersion.)

We computed adjusted R^2, a measure of goodness of fit that is resistant to bias from

small sample sizes or large numbers of covariates, as:

$$R^2_{\text{adj}} = 1 - \frac{D(y;\mu_{\text{full}}) + (k_{\text{full}} - k_{\text{null}})/v}{D(y;\mu_{\text{null}})}$$

This measure is analyzed in further detail in (Mittlböck & Heinzl, 2002).

Full counts of the number of channels included in the analysis in Figure 2-7 are included in Table 2.2.

Band and Epoch	LIP NPS	LIP PS	FEF NPS	FEF PS	lPFC NPS	lPFC PS	Total NPS	Total PS
Early Sample 8-50 Hz	9	11	25	22	52	28	86	81
Early Sample 50-100 Hz	16	25	35	22	60	49	111	96
Late Sample/Early Delay 8-50 Hz	35	30	32	64	90	105	157	199
Late Sample/Early Delay 50-100 Hz	22	39	41	44	78	93	141	176
Late Delay 8-50 Hz	28	24	31	43	81	82	140	149
Late Delay 50-100 Hz	10	26	38	40	85	81	133	147

Table 2.2: Number of electrodes analyzed in each region in 2-7. NPS = Not position-selective, PS = position-selective. Because only electrodes showing any significant power modulation in a given band and epoch were included in the analysis, total counts differ for each row.

2.3.8 Synchrony Analysis

To determine the relationship between synchrony and load (Figures 2-8, 2-9, 2-S2, 2-S3, and 2-S4), for each electrodes pair, time point, and frequency, we first computed surrogate coherence values with each individual trial removed:

$$s_i \equiv \text{coh}([x_1, x_2, \ldots, x_{i-1}, x_{i+1} \ldots, x_n], [y_1, y_2, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n])$$

where x and y are the wavelet or multitaper coefficient vectors for the two electrodes and $\mathrm{coh}(x, y)$ is the standard sample estimator of coherence:

$$\mathrm{coh}(x, y) = \frac{|x \cdot y|}{\|x\|\|y\|}$$

The value $\mathrm{coh}(x, y) - s_i$ provides a measure of the contribution of the ith trial to the total coherence. This measure has been used previously (Womelsdorf, Fries, Mitra, & Desimone, 2006; Hipp, Engel, & Siegel, 2011) and has recently been shown to achieve greater power than alternative approaches (Richter, Thompson, Bosman, & Fries, 2015). We then computed the correlation between these surrogates and load (denoted here as l) as $v_i = -\mathrm{corr}(\mathrm{coh}(x, y) - s, l) = -\mathrm{corr}(s, l)$, and the mean of these correlation values as $\tanh\left(\sum_i \tanh^{-1} v_i\right)$.

To compute synchrony statistics for individual load conditions (Figure 2-S4), we used the pairwise phase consistency, an estimator of the squared mean resultant length that is not biased by the number of trials (Vinck, van Wingerden, Womelsdorf, Fries, & Pennartz, 2010):

$$\begin{aligned}
\mathrm{PPC}(\theta) &= \frac{2}{N(N-1)} \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} (\cos\theta_j \cos\theta_k + \sin\theta_j \sin\theta_k) \\
&= \frac{1}{N(N-1)} \left[\left| \sum_{j=1}^{N} \exp(i\theta_j) \right|^2 - N \right] \\
&= 1 - \mathrm{var}(\cos\theta) - \mathrm{var}(\sin\theta)
\end{aligned}$$

where θ is the vector of angular differences between the coefficients of the two electrodes, i.e., $\theta_j = \arg(y_i) - \arg(x_j)$, $\exp(i\theta_j) = \frac{\bar{x}_j y_j}{|x_j||y_j|}$.

As in the power analyses, we determined confidence intervals and p-values in all synchrony analyses by bootstrapping over trials. We used 100 bootstraps for the time-frequency localized analysis in Figures 2-8 and 2-S2, and 1000 bootstraps for the bar graphs in Figures 2-9, 2-S3, and 2-S4 and p-values reported in the text.

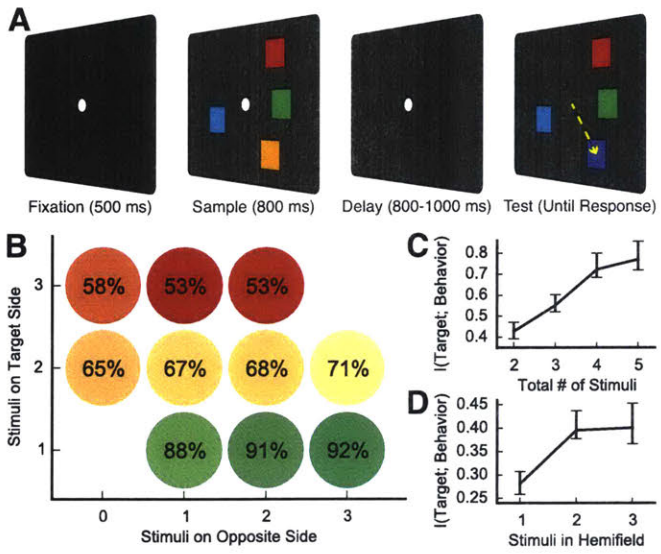


Figure 2-1: **A**: Change localization task. After fixating for 500 ms, animals saw an array of colored squares for 800 ms. These squares then disappeared, and subjects were required to maintain the colors of these squares in memory for a variable delay of 800-1000 ms. The array then reappeared with a change to the color of one square. The animal was rewarded for saccading to the changed square. **B**: Average behavioral performance according to the number of squares on the same side as the changed stimulus ("target side") and the number of squares on the opposite side. Performance depended on the number of squares on the target side, not the total number of squares. **C**: Mutual information between the location of the target stimulus and the animal's choice given the display for total loads 2 through 5. **D**: Mutual information between the location of the target stimulus and the animal's choice given the display for loads 1, 2, and 3 in the target hemifield. Error bars reflect 95% confidence intervals based on non-parametric bootstrapping across sessions.

2.4 Results

2.4.1 Monkeys Maintain Multiple Items in Working Memory

Our change localization task and behavioral results are described in detail in Buschman et al. (2011). Animals saw a sample array of 2 to 5 colored squares for 800 ms (Figure 2-1A). After an 800-1000 ms delay period, the array reappeared with a change to the color of one of the squares. Animals had to saccade to the changed square to receive a juice reward. The "target" was defined as the stimulus that changed between the two array presentations. The monkey was not cued to (and could not predict) which stimulus would be the target.

As Buschman et al. (2011) reported and others have confirmed (Matsushima & Tanaka, 2014), the right and left visual hemifields seemed to process stimuli separately. There were independent capacities on the right and left for the number of stimuli that could be processed/remembered: A stimulus in the same hemifield as the target degraded both the ability of the animal to detect its change and the neural information present about that stimulus. By contrast, a stimulus in the opposite hemifield had little or no effect. Indeed, task performance showed a strong dependence on the number of stimuli in the same hemifield as the target ($p < 10^{-10}$, ANOVA; Figure 2-1B) but no significant dependence on the number of stimuli in the opposite hemifield from the target ($p = 0.23$, ANOVA). To determine the animal's behavioral capacity, we measured mutual information between the animals' choices and the target position (Figure 2-1B). Mutual information between the target and response plateaued at 4 items in the display, consistent with reports of working memory capacity of 3 to 4 items in humans (Luck and Vogel 1997; Vogel et al. 2001). Because the behavioral analysis above suggested that the two hemifields were independent, we used linear regression to separate information in each hemifield (see Methods). Information increased when a second stimulus was added to the target hemifield ($p < 10^{-15}$, bootstrap Z-test; Figure 2-1D), but showed no change when a third stimulus was added ($p = 0.94$). Thus, behavioral capacity appears to saturate between 1 to 2 stimuli per hemifield.

2.4.2 Oscillatory Power Correlates with Stimulus Load

We tested the effects of contralateral stimulus load on LFP power using a generalized linear model. After removing evoked potentials, we regressed the instantaneous power at 22 frequencies spanning from 2.8 Hz to 107.6 Hz and 211 time points from -500 to 1600 ms after stimulus onset against the number of stimuli presented in the ipsilateral and contralateral hemifields (see Methods). We estimated the multiplicative contribution over the model intercept, which estimates the power in the absence of any stimuli. Figure 2-2A shows the percent power change for each contralateral stimulus load condition. Figure 2-2B plots the percent change in LFP power for each added contralateral stimulus, as estimated based on

observed power at loads one to three. The black polygons indicate time-frequency "zones" of significant changes in power with increasing load (see Methods). Figure 2-3 shows the same analysis for ipsilateral stimulus loads. We first consider the effects of contralateral loads. We saw effects of contralateral load in two broad bands, 8-50 Hz ("lower frequencies", including theta, alpha, beta, and low gamma) and 50-100 Hz ("higher frequencies" or high gamma). We chose these bands empirically based on the clear differentiation between them within the data. Although our "higher frequencies" are slightly higher in frequency than previous reports of gamma activity in early visual areas (Fries, Womelsdorf, Oostenveld, & Desimone, 2008), the peak of gamma modulation within visual areas differs according to stimulus contrast (Ray & Maunsell, 2010). Modulation in the 50-100 Hz band we investigate has been reported in human MEG studies (Siegel, Donner, Oostenveld, Fries, & Engel, 2008; Hipp et al., 2011) as well as other monkey studies of oscillations in frontal regions (Lara & Wallis, 2014; Lundqvist et al., 2016). We will first summarize the main effects. Then, we will discuss their temporal dynamics.

During sample array presentation, higher frequency LFP power increased with increased contralateral stimulus load (positive correlations, warm colors) while lower frequency LFP power decreased with increased contralateral stimulus load (negative correlations, cool colors). During the memory delay (especially late in the delay), there was no effect of contralateral load on higher frequencies and the effects at lower frequencies inverted (Figure 2-2).

Figure 2-3 shows the effect of ipsilateral stimulus load on LFP power. Like contralateral load, ipsilateral load inversely correlated with lower frequency LFP power during stimulus presentation. But in contrast to contralateral load, effects of ipsilateral load at higher frequencies were weaker. Additionally, the positive correlation between memory delay lower frequency lPFC power and contralateral load was not observed for ipsilateral loads. Thus, the independence between the two visual hemifields seen in behavior was reflected in positive, but not the negative, correlations between load and power. As Figures 2-2 and 2-3 illustrate, the changes in LFP with stimulus load was more complex than this summary; there were

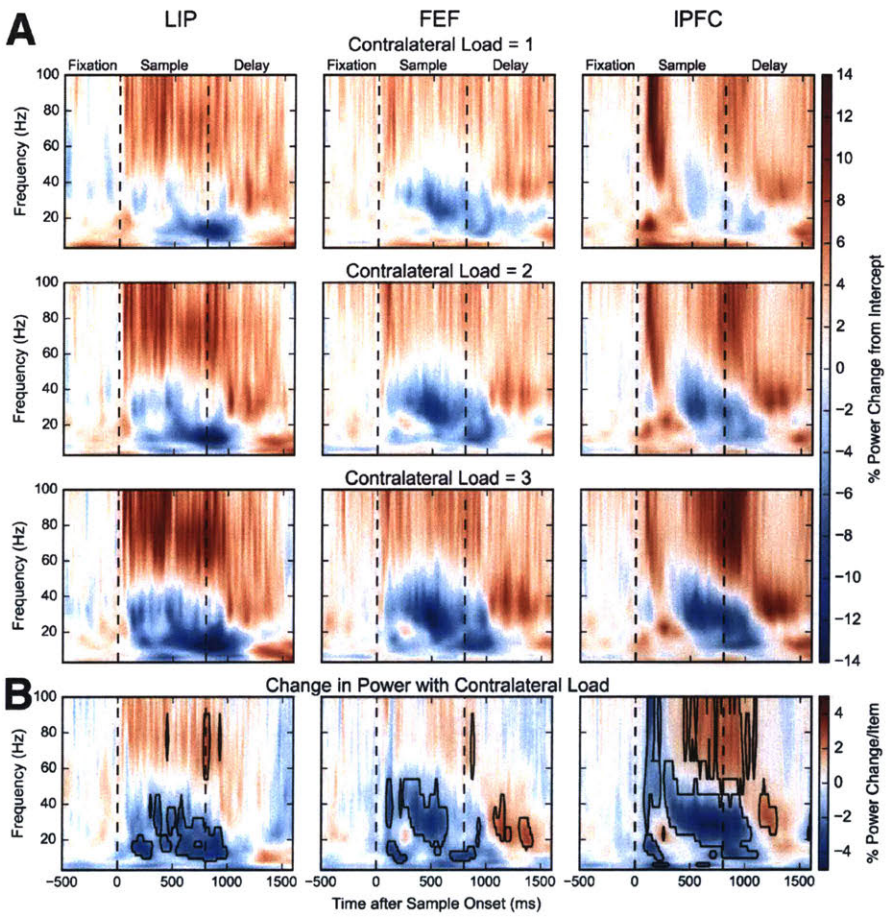


Figure 2-2: **A**, Percent power change for contralateral loads 1, 2, and 3 relative to the model intercept across frequencies and time. First dashed line indicates time of sample onset. Second dashed line indicates time of sample offset. **B**, Percent power change per contralateral stimulus. Boxes indicate significant modulations (bootstrap Z-test, $p < 0.05$, Holm corrected for 22 frequencies $\times$ 211 time points).

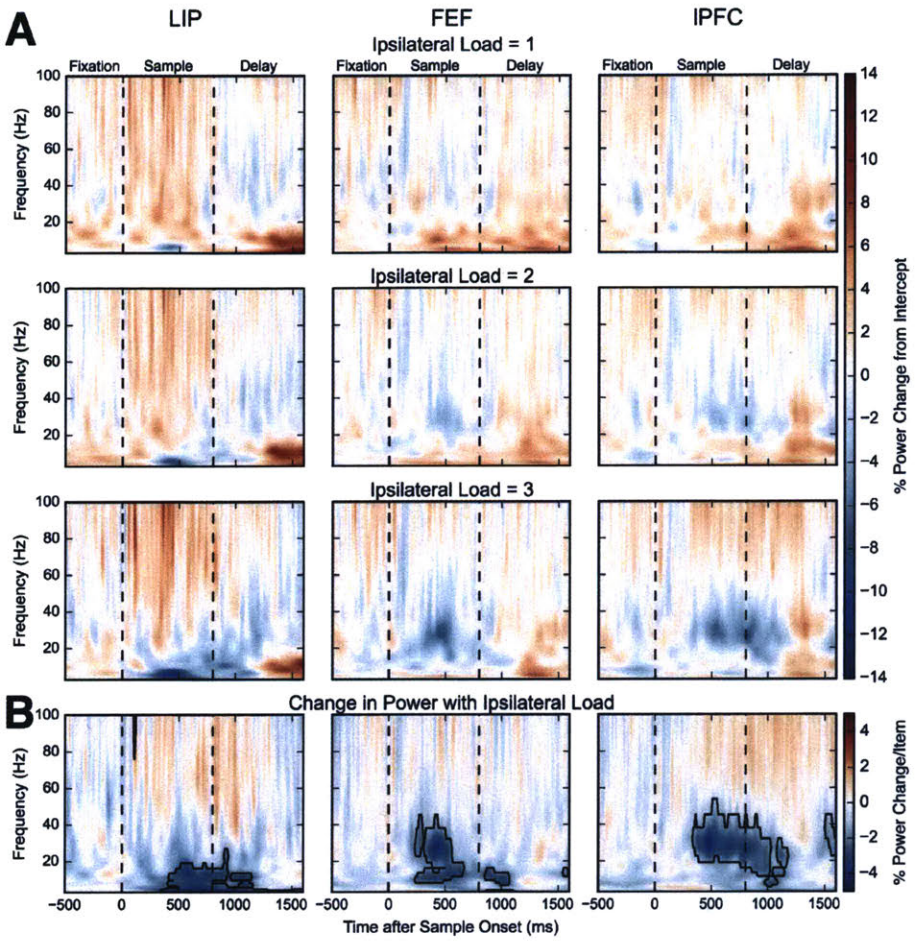


Figure 2-3: **A**, Percent power change for ipsilateral loads 1, 2, and 3 relative to the model intercept across frequencies and time. First dashed line indicates time of sample onset. Second dashed line indicates time of sample offset. **B**, Percent power change per ipsilateral stimulus. Boxes indicate significant modulations (bootstrap Z-test, $p < 0.05$, Holm corrected for 22 frequencies $\times$ 211 time points).

differences in the temporal dynamics in effects between brain areas. Next, we quantify these effects and consider them in more detail.

2.4.3 Differences in Effects of Stimulus Load by Time and Frequency Band

As Figures 2-2 and 2-3 indicate, the main effects of stimulus load were seen across two broad frequency bands (lower: 8-50 Hz and higher: 50-100 Hz). When we separated the lower

frequencies into standard frequency bands (theta, 4-8 Hz; alpha, 8-12 Hz; beta, 16-30 Hz; low gamma, 30-50 Hz), the pattern of effects in the narrow bands mirrored the broadband effects (i.e., negative correlations in theta through low gamma, positive correlations in high gamma, see Figure 2-S1). Thus, for simplicity we report statistics computed across the broad bands using the multitaper method (see Methods).

Figures 2-2 and 2-3 suggest that different load effects grouped themselves into three distinct time periods, especially in the lPFC. There was an early sample epoch (0-400 ms after sample onset), a late sample/early delay epoch (400 – 1000 ms after sample onset) and a late delay epoch (1000-1800 ms after sample onset). We therefore computed the power across each of these epochs separately. Figure 2-4 plots the average percent change in LFP power per added stimulus for the two broad frequency bands and in each of the three epochs for each brain area. The asterisks indicate a significant change in power with increasing load. The hatched bars indicate when a given measure in a given brain area showed a significant difference between the effects of contralateral vs. ipsilateral load. Next, we consider the effects in each epoch.

Early Sample Epoch (Figure 2-4A)

All three brain areas showed a decrease in low frequency power with increased contralateral stimulus load (LIP: -1.4%/item, $p < 10^{-6}$; FEF: -1.2%/item, $p < 10^{-8}$; lPFC: -0.9%/item, $p < 10^{-6}$). The effects of contralateral load on higher frequency power were mixed. LIP showed a significant positive correlation (0.8%/item, $p = 0.004$) and lPFC showed a significant negative correlation with contralateral load (-0.8%/item, $p < 10^{-6}$). For ipsilateral loads, there were numerically negative correlations with lower frequencies, but only the FEF and lPFC showed a significant negative correlation with lower frequency power (FEF: -1.1%/item, $p < 10^{-6}$; lPFC: -0.5%/item, $p = 0.03$). The differences in correlations for contralateral vs. ipsilateral stimulus loads did not reach significance for the lower frequencies. (No bars corresponding to lower frequencies in Figure 2-4A are hatched, indicating no difference between contralateral and ipsilateral loads for each area.) In contrast to contralateral

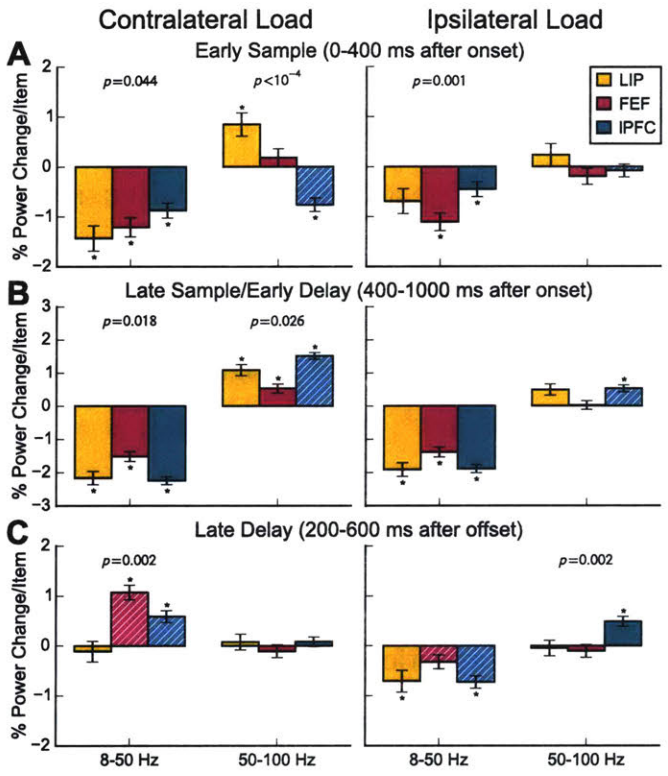


Figure 2-4: Percent power change per contralateral (left) and ipsilateral (right) item by region, grouped by lower frequencies (left bar group) and higher frequencies (right bar group) during the early sample (A), late sample/early delay (B), and late delay (C). Error bars are standard error of the mean. Asterisks indicate significant differences (bootstrap Z-test, $p < 0.05$, Holm corrected for 2 bands $\times$ 3 epochs $\times$ 3 regions). White hatching indicates significant differences in modulation by ipsilateral and contralateral load (bootstrap Z-test, $p < 0.05$, Holm corrected). p-values above bars indicate significant differences between regions (F-test, $p < 0.05$).

loads, none of the areas showed a significant correlation between higher frequencies and ipsilateral load. This difference between the effects of contralateral and ipsilateral load on higher frequencies was significant for the lPFC ($p = 0.007$; hatched bars, Figure 2-4A). In this way, the higher frequency power was similar to our previously reported single-neuron results (Buschman et al. 2011).

Late Sample/Early Delay Epoch (Figure 2-4B)

Contralateral stimulus load had different effects on lower versus higher frequencies in all three areas. All three areas showed a significant negative correlation with contralateral load at lower frequencies (LIP: -2.2%/item, $p < 10^{-25}$; FEF: -1.5%/item, $p < 10^{-22}$; lPFC: -2.2%/item, $p < 10^{-79}$; see Figure 2-4B) and a significant positive correlation at higher frequencies (LIP: 1.1%/item, $p < 10^{-8}$; FEF: 0.5%/item, $p = 0.001$; lPFC: 1.5%/item, $p < 10^{-45}$). As in the early sample epoch, lower frequency power decreased with ipsilateral load in all three areas (LIP: -1.9%/item, $p < 10^{-18}$; FEF: -1.4%/item, $p < 10^{-19}$; lPFC: -1.9%/item, $p < 10^{-50}$), with weak or no effect on higher frequencies. Only the lPFC showed a small, but significant positive correlation between higher frequency power and ipsilateral load (0.5%/item, $p < 10^{-5}$) and it was significantly weaker than the correlation between lPFC higher frequency power and contralateral load (hatched bar in Figure 2-4B; $p < 10^{-9}$).

Late Delay Epoch (Figure 2-4C)

Later in the memory delay, the effects of contralateral load on lower frequency power in the FEF and lPFC reversed relative to earlier in the trial. They showed a significant positive (as opposed to negative) correlation between lower frequency power and contralateral load (FEF: 1.1%/item, $p < 10^{-11}$; lPFC: 0.6%/item, $p < 10^{-4}$). By contrast, there was no effect of contralateral load on higher frequency power. Increased ipsilateral stimulus load continued to produce negative correlations with lower frequency power in LIP (-0.7%/item, $p = 0.02$) and lPFC (-0.7%/item, $p < 10^{-7}$). This difference between the effects of contralateral and ipsilateral load at lower frequencies was significant for the FEF and lPFC (hatched bars,

Figure 2-4C; FEF: $p < 10^{-9}$; lPFC: $p < 10^{-11}$). Ipsilateral load effects on higher frequency power remained weak; only the lPFC showed a positive correlation (0.5%/item, $p < 0.0001$).

2.4.4 Do Load Effects Saturate at Behavioral Capacity?

Above, we cataloged whether increased stimulus load increased or decreased LFP power. These effects could reflect a strictly monotonic relationship (every added stimulus changes LFP power to a certain degree). Alternatively, there could be a step-like, or threshold, relationship. For example, power could have been constant below a specific load and then increased above it. To test this, we computed the average change from baseline (i.e., no stimulation) for each contralateral and ipsilateral load from one to three stimuli. This is plotted in Figure 2-5 for each area and for contralateral and ipsilateral loads. As can be seen, the relationship between stimulus load and power seems mostly monotonic. Each added stimulus produces a similar degree of change in oscillatory power. Note that, even though the animals' performance dropped off beyond two stimuli, adding a third stimulus to the load resulted in further changes in power. The asterisks in Figure 2-5 indicate which brain areas showed a strictly monotonic relationship between power and load, i.e., the smallest observed difference in power of any pair of load conditions was larger than would be expected by chance (see Methods). In most bands, epochs, and regions where we observed significant modulation of power by load, the effect was strictly monotonic. This was true for all observed low frequency power decreases with contralateral or ipsilateral load in the early sample and late sample/early delay periods, as well as high frequency power increases with contralateral load in LIP during the early sample and in LIP and lPFC during the late sample and early delay. Crucially, in FEF and lPFC, strictly monotonic and seemingly linear effects were present even in the late delay period.

2.4.5 Relationship Between LFP Position and Load Information

In addition to information about stimulus load, LFPs also carried information about stimulus positions. Similar to above, we fit generalized linear models to each electrode, band, and

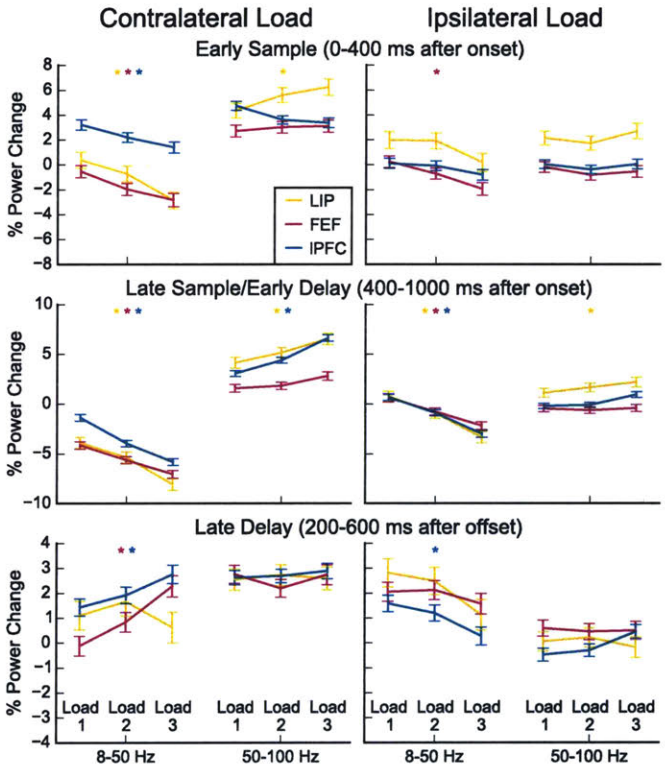


Figure 2-5: Percent power change for contralateral loads (left) and ipsilateral loads (right) 1, 2, and 3 relative to load 0, for epochs and frequency bands. Asterisks indicate significance of all pairwise differences for the band, region, and epoch (permutation test, $p < 0.05$, Holm corrected for 2 bands $\times$ 3 epochs $\times$ 3 regions).

epoch incorporating either load alone or both load and stimulus positions in either the contralateral or ipsilateral hemifields. Based on the difference in these two models, we computed adjusted R^2, a debiased measure of the additional information captured by the model that included stimulus positions similar to ω^2 in linear ANOVA (see Supplementary Methods).

In all bands, epochs, and regions, significant information about stimulus position was present in LFP power (all $p < 0.0003$, non-parametric bootstrap test, Holm-corrected; Figure 2-6). Additionally, a small amount of information about ipsilaterally presented stimuli was present in low frequency power in the late sample/early delay period and late delay periods in all regions and at high frequency power in FEF and lPFC. However, the amount of information about contralaterally presented stimuli was substantially greater in all bands, frequencies, and epochs (all $p < 0.005$, paired non-parametric bootstrap test, Holm-corrected).

The modulation of power by load reported above could reflect either position-specific or position-invariant effects. If power in a given electrode increased or decreased when a stimulus was presented in a specific location, then power might also increase or decrease with load, since any given stimulus is more likely to be present at higher loads than at lower loads. Alternatively, the load effects may reflect global changes in power, i.e., even electrodes that do not carry position information might nonetheless show modulation by load. To distinguish these possibilities, for each frequency band and epoch, we separately computed the average percent power change per contralateral stimulus for electrodes with and without significant effects of contralateral stimulus position in that epoch and band (F-test, $p < 0.05$). To minimize confounds from non-position selective electrodes with low statistical power, we excluded electrodes where neither load nor stimulus position explained any significant variation. Because previous analyses showed similar trends across recorded regions, we pooled electrodes across regions to increase statistical power. Applying the analysis to only frontal (FEF and lPFC) electrodes yielded an identical pattern of significance, as did a test for a significant main effect of position selectivity in a 2×3 ANOVA. There were no significant interactions between the effect of position selectivity and region

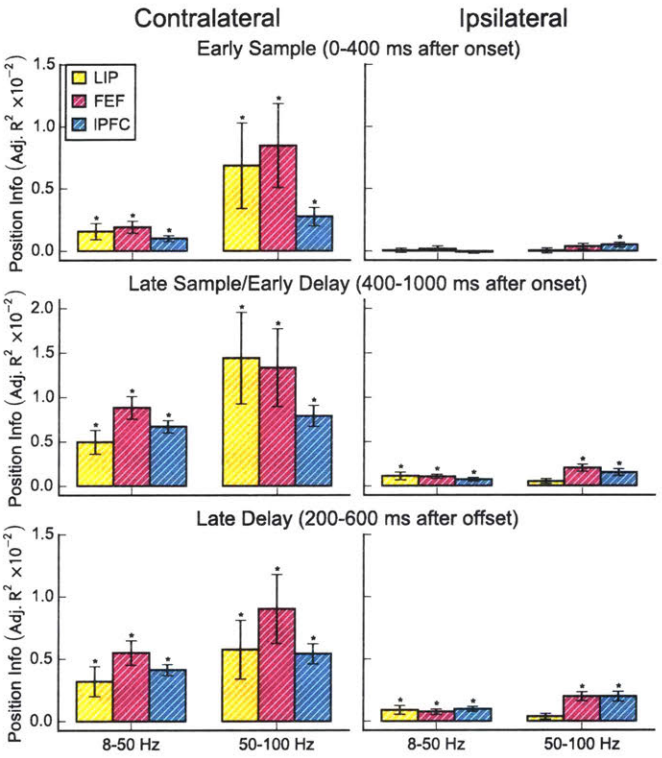


Figure 2-6: Comparison of position information (adjusted R^2) for contralateral and ipsilateral stimuli. Error bars are standard error of the mean. Asterisks indicate significant information (non-parametric bootstrap test, $p < 0.05$, Holm corrected for 2 bands $\times$ 3 epochs $\times$ 3 regions). White hatching indicates significant differences in modulation by ipsilateral and contralateral load (non-parametric paired bootstrap test, $p < 0.05$, Holm corrected).

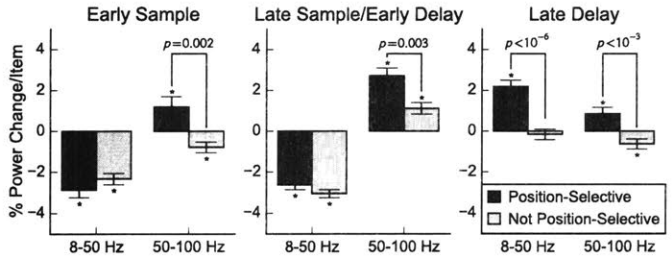


Figure 2-7: Percent power change per contralateral item for position-selective and non-position-selective electrodes. Asterisks indicate significant modulation by load (one sample t-test, $p < 0.05$, Holm corrected for 2 bands $\times$ 3 epochs). p-values above bars indicate significant differences between position-selective and non-position-selective electrodes (unequal variance t-test, $p < 0.05$, Holm corrected).

(all $p > 0.15$, Holm-corrected F-test). The results of this analysis are shown in Figure 2-7.

In the early sample and late sample/early delay period, lower frequency power decreased with contralateral load in both position-selective and non-selective electrodes (all $p < 10^{-9}$, t-test), but the strength of the modulation did not differ (early sample: $p = 0.34$ [0.26 uncorrected]; late sample: p = 0.34 [0.17 uncorrected]; unequal variance t-test with Holm correction). Differences were present at higher frequencies. In the early sample period, power in position-selective electrodes increased with contralateral load (1.2%/item, $p = 0.04$, one-sample t-test with Holm correction), whereas power in non-position-selective electrodes decreased with load (-0.8%/item, $p = 0.01$; difference: $p = 0.002$, unequal variance t-test with Holm correction).

In the late sample/early delay period, higher frequency power in both position-selective and non-position-selective electrodes increased with load (position-selective: 2.7%/item, $p < 10^{-10}$; non-position-selective: 1.1%/item, $p = 0.0004$), but position-selective electrodes showed a stronger average modulation ($p = 0.003$). Thus, while the mean decrease in lower frequency power with load in the sample period is independent of position selectivity, the mean increase in higher frequency power with load appears to be driven largely by position-selective electrodes.

In the late delay period, lower frequency power in position-selective electrodes increased with load (2.2%/item, $p < 10^{-9}$). Non-position-selective electrodes showed no average mod-

ulation (-0.16%/item, p = 0.5; difference: $p < 10^{-6}$). Since these electrodes were selected on the basis of the presence of power or load effects, we conjecture that the individual channel effects average to zero over the recorded population. Thus, like the increase in higher frequency power during the sample period, the mean increase in lower frequency power with load in the late delay period appears to be due to position-selective electrodes. Higher frequency power increased with load for position-selective electrodes (0.9%/item, $p = 0.02$) and decreased with load for non-position-selective electrodes (-0.6%/item, $p = 0.03$; difference: $p < 0.001$). This relationship between load and stimulus position effects suggests that the effects of load mostly occur in neuron populations that process bottom-up information about the stimuli (see Discussion).

2.4.6 Effects of Stimulus Load on LFP Synchrony

The above analyses focused on changes in oscillatory power with stimulus load. We next examine whether stimulus load affected the synchrony of LFP signals between electrodes within and across brain areas. We computed total coherence values across all trials, and constructed single trial surrogate coherence values for each trial as the difference between these total coherence values and coherence values based on all trials except the trial of interest. We then measured the correlation between LFP synchrony and load as the correlation between these single trial coherence surrogates and contralateral or ipsilateral load in that trial (see Methods).

Figure 2-8 shows the correlation between LFP synchrony and contralateral stimulus load as a function of frequency and time during the trial (see Methods). There was little, if any, change in synchrony with ipsilateral load (Figures 2-S2 and 2-S3). Figure 2-9 shows the effects of load on synchrony for the standard frequency bands (theta, alpha, beta, and low and high gamma). Here, we summarize the significant effects using the same higher/lower frequency classifications we used for LFP power.

During sample presentation and shortly after, increases in contralateral load increased higher frequency LFP synchrony within LIP (early sample: $r = 0.017$, $p = 0.0002$, bootstrap

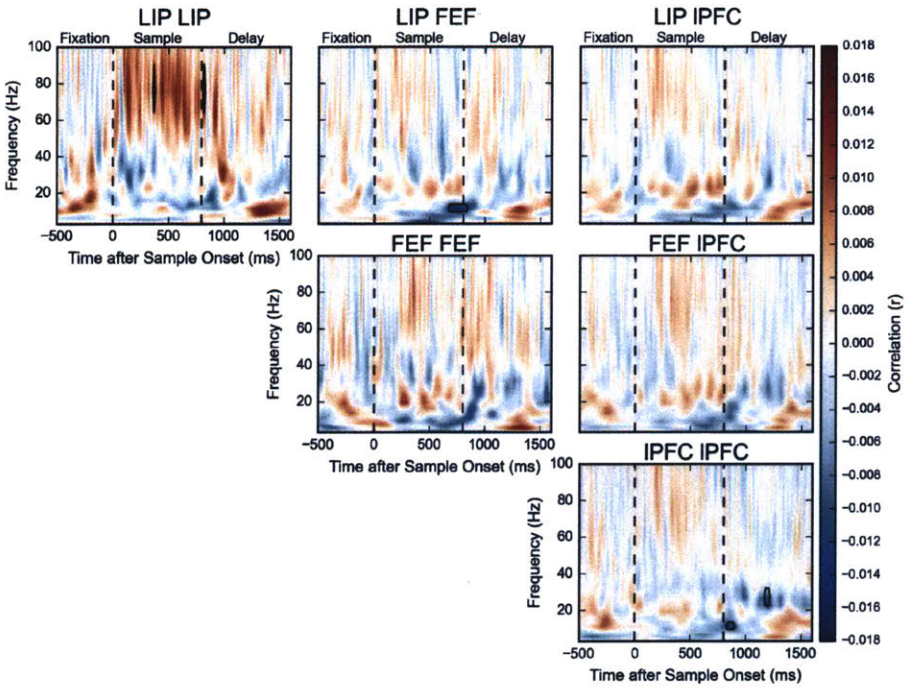


Figure 2-8: Correlation of single trial coherence surrogates with contralateral load. Boxes indicate significant modulations (bootstrap Z-test, $p < 0.05$, Holm corrected for 22 frequencies $\times$ 211 time points). The same analysis for ipsilateral load is shown in Figure 2-S2.

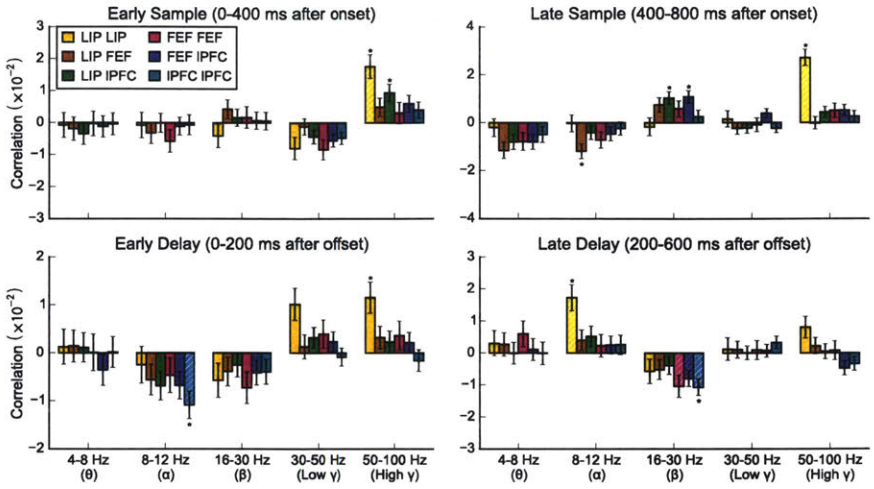


Figure 2-9: Correlation of single trial coherence surrogates with contralateral load for frequency bands and epochs. Asterisks indicate significant differences (bootstrap Z-test, $p < 0.05$, Holm corrected for 5 bands $\times$ 4 epochs $\times$ 6 region pairs). White hatching indicates significant differences in modulation by ipsilateral and contralateral load (bootstrap Z-test, $p < 0.05$, Holm corrected). The same analysis for ipsilateral load is shown in Figure 2-S3.

Z-test, Holm-corrected; late sample: $r = 0.023$, $p < 10^{-12}$; early delay: $r = 0.012$, $p = 0.04$; Figures 2-8 and 2-9) and between LIP and lPFC (only in the early sample, $r = 0.009$, $p = 0.049$). In the late sample epoch, increased contralateral load significantly increased beta synchrony between LIP and lPFC ($r = 0.010$, $p = 0.02$) and between the FEF and lPFC ($r = 0.011$, $p = 0.004$). (While these cross-region effects did not achieve significance in the time-localized analysis in Figure 2-8, they are apparent in the epoched analysis in Figure 2-9.) In the late delay, beta synchrony within the lPFC instead decreased with contralateral load ($r = -0.011$, $p = 0.004$). Additionally, there were significant decreases in alpha synchrony with contralateral load between LIP and FEF in the late sample ($r = -0.012$, $p = 0.02$) and within the lPFC in the early delay ($r = -0.011$, $p = 0.02$), and an increase in alpha synchrony with contralateral load within LIP in the late delay ($r = 0.017$, $p = 0.002$). As noted above, there were few changes in synchrony with ipsilateral load. There was only a decrease in alpha synchrony in the late sample period ($r = -0.012$, $p = 0.009$, see Figures 2-S2 and 2-S3).

Synchrony effects were generally fewer and smaller than the effects of load on LFP power. Given the small effect sizes, it is possible that these synchrony effects are not computationally relevant. Nonetheless, we were able to detect that the effects of contralateral load on LFP synchrony in the sample period had a monotonically increasing relationship with contralateral load. Figure 2-S4 shows the difference in pairwise phase consistency between trials with contralateral loads 1, 2, and 3 and zero contralateral load trials for each region pair, band, and epoch. We determined significance using the same permutation test as in the test for monotonicity of power changes above. Synchrony within LIP increased monotonically with load in the higher frequencies in the early and late sample periods. Synchrony between LIP and lPFC increased monotonically with contralateral load in the beta band in the late sample period. We did not see significant monotonic effects in other bands after the multiple comparison correction.

2.5 Discussion

We tested the effects of different stimulus loads in three brain areas known to be important for visual attention and working memory, the lateral prefrontal cortex (lPFC), the frontal eye fields (FEF) and the lateral intraparietal area (LIP). Previously, we reported the single neuron correlates of stimulus load (Buschman et al., 2011). Here, we report the effects of load on oscillatory dynamics (power and synchrony) of local field potentials (LFPs).

During stimulus presentation, there were decreases in lower frequency (8-50 Hz) LFP power with increases in contralateral and ipsilateral stimulus load. However, higher frequency (50-100 Hz) LFP power increased only with contralateral stimulus load. Contralateral load also (briefly) increased lower frequency (16-50 Hz, encompassing beta and low gamma) power in the middle of the memory delay. This is relevant because our monkeys showed behavioral evidence of separate visual working memory capacities in each hemifield. This predicts that the neural effects of load tied most directly to behavior would be limited to the contralateral field. Synchrony measures are naturally noisier than measures of power but the effects of

load on synchrony generally matched effects on power. The exceptions were in the beta band where there was an inverse relationship between stimulus load and synchrony. Beta synchrony (16-30 Hz) increased with stimulus load during the sample presentation (when beta power decreased) and decreased during the delay (when beta power increased). Finally, we found that even after the monkeys' behavioral capacity had been reached, increases in stimulus load continued to affect oscillatory power.

Our results coincide with previous reports of changes in LFP power in multiple item working memory tasks. Lara and Wallis (2014) also found that in the lPFC, high gamma power increased and beta power decreased during stimulus presentation, while beta power increased during a memory delay. While they did not explicitly compare power in one-item and two-item trials, the effects appear to be stronger in the two-item trials, consistent with our findings that power in these bands and epochs scales with load. S. Palva et al. (2011) analyzed MEG and EEG power during a multiple item working memory task in humans. They reported increases in delay period beta and low gamma power with load that did not saturate at behavioral capacity. However, the same authors report that power increases were associated with strengthened inter-areal synchrony (J. M. Palva et al., 2010), which we also observed during sample presentation, but not in the memory delay. Mitchell and Cusack (2011) showed bilateral decreases in induced alpha power immediately following sample presentation in human MEG data, consistent with our findings during the sample presentation.

Increases in high frequency power accompanied by decreases in lower frequency power have also been reported from a variety of visual attention and perception tasks. Monkey studies of V4 and FEF LFPs have shown that attention towards a stimulus increases higher frequency power and decreases lower frequency power (Fries et al., 2008; Gregoriou, Gotts, Zhou, & Desimone, 2009; Gregoriou, Rossi, Ungerleider, & Desimone, 2014). Human EEG and MEG studies report similar results across a wide range of cortical regions (Siegel et al., 2008; Hipp et al., 2011). Our findings add important details. We found a dissociation between the effects of load on lower versus higher frequency power and their relationship

to behavior. During sample presentation, lower frequency power was modulated by both contralateral and ipsilateral load, whereas higher frequency power was modulated more by contralateral load. This reflects the strong contralateral bias in visual cortical processing. It suggests that the effects of load on higher frequencies are more strongly associated with bottom-up processing of visual stimulus information per se. Indeed, we found stronger higher frequency power changes in electrodes that showed selectivity for stimulus location (whereas lower frequency power changes were similar in all electrodes.).

This may be due to previously observed associations between lower and higher frequency oscillations and top-down and bottom-up cortical processing, respectively (Engel et al., 2001; Buschman & Miller, 2007; Engel & Fries, 2010; Arnal & Giraud, 2012; Bastos et al., 2015). Lower frequency (beta) oscillations have been linked to maintaining the existing cognitive set (Engel & Fries, 2010; Buschman et al., 2012) and may help to stabilize working memory representations against disruption during memory delays (Pereira & Wang, 2014). Our data support these hypotheses. During sample presentation, higher frequency power increased with contralateral load, especially in electrodes with bottom-up information about stimulus location. The difference in modulation between position-selective and non-position-selective electrodes suggests that higher frequency oscillations reflect bottom-up input from sensory areas. However, since non-position-selective electrodes were also modulated by contralateral load during the late sample/early delay period, albeit more weakly, these oscillations could additionally reflect top-down modulation of cortical areas processing contralateral stimuli. Lower frequency power decreased with load during sample presentation and in all electrodes. This suggests a more global state change such as a broader focusing of attention across more locations. By contrast, during the late delay, lower frequency power instead increased only with contralateral load and only in electrodes with information about stimulus location. Thus, during memory maintenance, beta oscillations may stabilize the working memories in the circuits that carry information about the stimuli.

Note that while beta power decreased during stimulus array presentation, beta synchrony increased. Beta synchrony has been linked to shifts of attention between multiple stimuli

(Buschman & Miller, 2009). Thus, the increase in beta synchrony with stimulus load may reflect an increased number of attentional shifts. One possible source for beta signals is the pulvinar, which projects to both prefrontal and parietal cortex (Asanuma, Andersen, & Cowan, 1985), contains units with both ipsilateral and contralateral receptive fields (Bender, 1981), and has recently been shown to modulate low frequency oscillations and synchrony in extrastriate visual cortex during attention (Saalmann, Pinsk, Wang, Li, & Kastner, 2012).

Our results put some constraints on models of how cognitive capacity arises. According to "slot" models, capacity is limited by an individual's specific number of memory slots. Once they are filled, capacity is reached (Luck & Vogel, 1997, 2013; Vogel et al., 2001; Ma et al., 2014). Any further increase in stimulus load should have no effect on neural activity; once all the slots are filled no more information can be encoded. By contrast, in flexible resource models information is a like a pool. Increasing stimulus load uses more and more of this pool. Once the pool becomes too thin, behavior can no longer be supported and effective capacity is reached, but increasing load beyond behavioral capacity will continue to draw from the pool and thus continue to affect neural activity (Bays & Husain, 2008; Luck & Vogel, 2013; Ma et al., 2014). Our subjects had a behavioral capacity of between one and two in each hemifield. However, we observed increases in local field potential power with stimulus load between two and three stimuli, indicating that the animals processed information about stimuli above behavioral capacity. While the absence of saturation at capacity during stimulus presentation might relate to purely visual processes, the absence of saturation during the late delay period is more surprising and more difficult to explain with a slot model. Nonetheless, it is possible that the power increase reflects maintenance of more spatial locations, but the capacity bottleneck arises in maintaining color per se. Thus, our results are consistent with either a resource model or a modified slot model in which information is maintained about positions of unremembered stimuli, but a fixed number of slots are available for object identity information.

In sum, we found increases in high gamma oscillations with increased stimulus load may reflect changes in feed-forward (bottom-up) sensory processing. Decreases in lower frequency

oscillations may instead reflect top-down processes such as the allocation of attention and working memory maintenance.

2.6 Contributions and Acknowledgements

I analyzed the data presented in this chapter and prepared the initial manuscript. Timothy J. Buschman and Earl K. Miller designed the experiment. Timothy J. Buschman collected the data. All authors commented on the manuscript.

We thank A. M. Bastos, S. L. Brincat, R. Loonis, M. Lundqvist, and A. S. Widge for comments on the manuscript. This work was supported by an NIMH Conte Center grant, The Picower Foundation, and an NSF Graduate Research Fellowship.

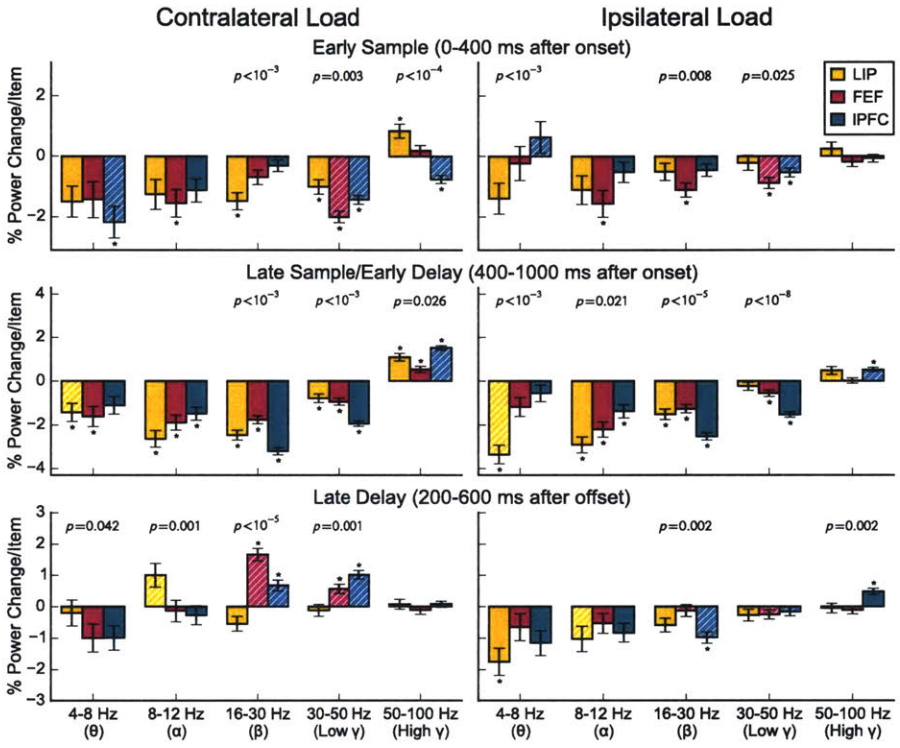


Figure 2-S1: Percent power change per contralateral (left) and ipsilateral (right) item by region for an extended set of frequency bands. Error bars are standard error of the mean. Asterisks indicate significant differences (bootstrap Z-test, $p < 0.05$, Holm corrected for 2 bands $\times$ 3 epochs $\times$ 3 regions). White hatching indicates significant differences in modulation by ipsilateral and contralateral load (bootstrap Z-test, $p < 0.05$, Holm corrected). p-values above bars indicate significant differences between regions (F-test, $p < 0.05$).

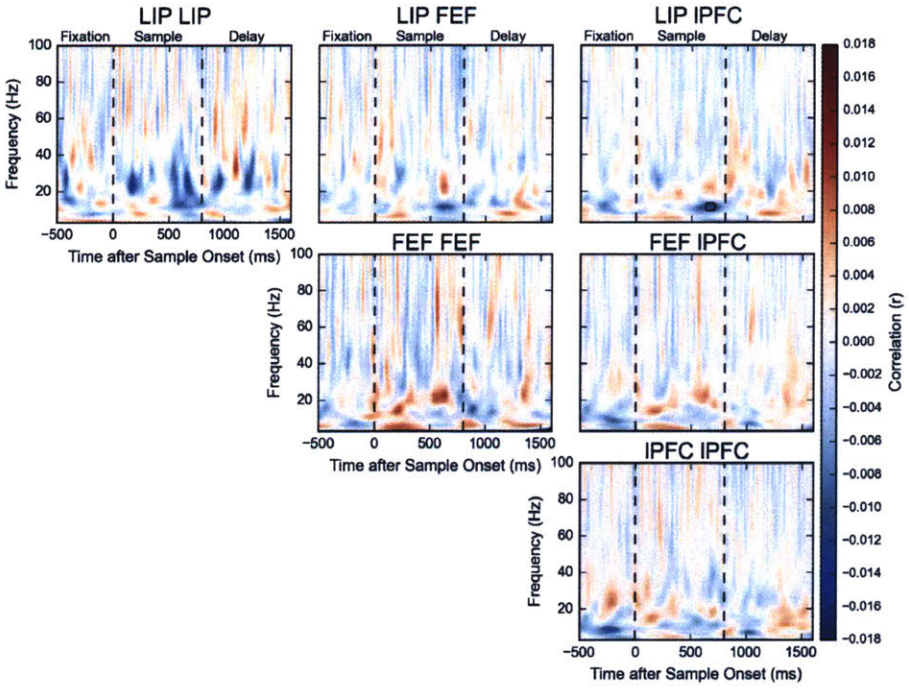


Figure 2-S2: Correlation of single trial coherence surrogates with ipsilateral load. Boxes indicate significant modulations (bootstrap Z-test, $p < 0.05$, Holm corrected for 22 frequencies $\times$ 211 time points). See Figure 2-8 for correlation with contralateral load.

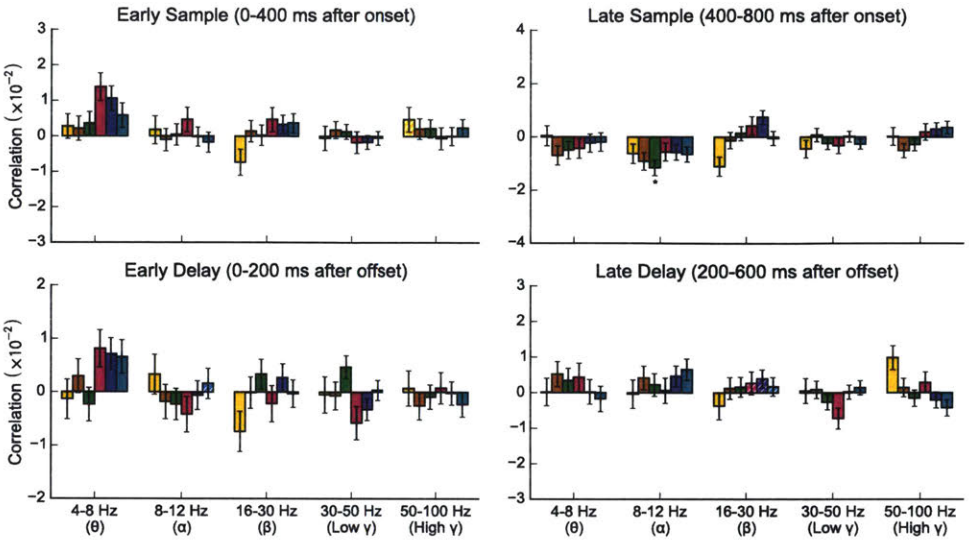


Figure 2-S3: Correlation of single trial coherence surrogates with ipsilateral load for frequency bands and epochs. Asterisks indicate significant differences (bootstrap Z-test, $p < 0.05$, Holm corrected for 5 bands $\times$ 4 epochs $\times$ 6 region pairs). White hatching indicates significant differences in modulation by ipsilateral and contralateral load (bootstrap Z-test, p$<$ 0.05, Holm corrected). See Figure 2-9 for correlation with contralateral load.

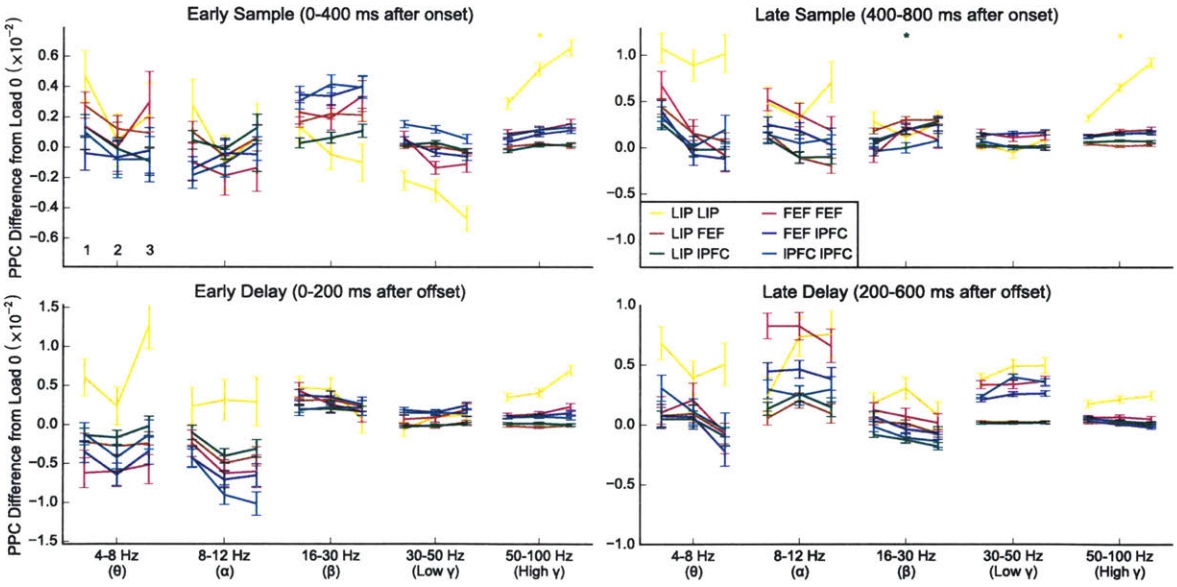


Figure 2-S4: Differences in pairwise phase consistency relative to no contralateral stimulation condition for contralateral loads 1, 2, and 3 (left to right), for epochs and frequency bands. Asterisks indicate significance of all pairwise differences for the band, region pair, and epoch (permutation test, $p < 0.05$, Holm corrected for 5 bands $\times$ 4 epochs $\times$ 6 region pairs).

Chapter 3

Neural Correlates of Multiple-Item Visual Short-Term Memory in Prefrontal, Parietal, and Temporal Cortex

3.1 Abstract

The richness of visual experience far exceeds our ability to remember what we have seen. However, it is unclear what neural mechanisms give rise to these limits to visual short-term memory capacity. Here, we recorded from the lateral intraparietal area (LIP), prefrontal cortex (PFC), and inferotemporal cortex (IT) during a multiple item change localization task as a monkey maintained multiple colored squares in visual short-term memory. Although activity in all regions reflected both stimulus positions and colors, LIP and PFC preferentially represented position, while IT preferentially represented color. At sample presentation, PFC conveyed more color information as the number of stimuli (load) increased, while IT conveyed the same amount of color information regardless of load. At test presentation, both PFC and IT conveyed the location of the changed square, with similar latencies. Although IT

delay period activity was barely detectable, activity at test signaled differences between sample and test displays, and the strength of this representation correlated with behavioral performance. Our results suggest that high-level visual cortex is more strongly capacity limited than prefrontal or parietal cortices, and imply a role of passive memory signals from visual regions in some visual short-term memory tasks.

3.2 Introduction

Most electrophysiological studies of visual short-term memory examine activity related to maintenance of a single object. However, in our everyday lives, objects are rarely present in isolation; our visual field typically contains dozens of objects. Our ability to remember the characteristics of these objects is limited, even over a brief delay (Simons & Rensink, 2005). When humans must remember more than 3 to 5 ordinary objects simultaneously, performance drops substantially (Luck & Vogel, 1997). Many studies have investigated how these capacity limits are manifested in behavior. Although some authors have proposed that items held in visual short-term memory are stored in a fixed number of discrete slots (Luck & Vogel, 1997; Awh et al., 2007; Zhang & Luck, 2008; Fukuda, Awh, & Vogel, 2010), the balance of evidence favors a model in which the total amount of representable information is fixed, but this "shared resource" is distributed flexibly across items (Bays & Husain, 2008; Bays et al., 2009; van den Berg et al., 2012; van den Berg, Awh, & Ma, 2014; Ma et al., 2014; Schneegans & Bays, 2016; Brady et al., 2016). Most of our knowledge of the neural mechanisms behind these capacity limits comes from the measures that reflect average activity of hundreds of thousands of neurons. Such studies typically implicate posterior parietal cortex, where functional magnetic resonance imaging (fMRI) activity and evoked potentials measured by electroencephalography (EEG) correlate with the number of items to be remembered (stimulus load), saturate at behavioral capacity, and reflect individual differences in task performance (Vogel & Machizawa, 2004; Todd & Marois, 2004; Vogel et al., 2005; Todd & Marois, 2005).

However, these bulk measures of neural activity measure only how average activity at a macroscopic scale changes with memory demand. They provide little insight into how the brain's stimulus representations change when multiple stimuli must be processed simultaneously. As the number of items to be remembered increases, stimulus representations could be severely degraded even if average activity does not change. Indeed, single neuron studies provide strong evidence that simultaneous presentation of multiple objects impairs stimulus representation in high-level visual areas (Sato, 1989; E. K. Miller, Gochin, & Gross, 1993; Rolls & Tovee, 1995; Chelazzi, Duncan, Miller, & Desimone, 1998; Missal, Vogels, Li, & Orban, 1999; Zoccolan, Cox, & DiCarlo, 2005). However, few single neuron studies have investigated representation of simultaneously presented stimuli in a visual short-term memory task. Buschman et al. (2011) reported that representations in the lateral intraparietal area (LIP) and prefrontal cortex (PFC) were degraded when subjects maintained multiple items (colored squares) in visual short-term memory, but Lara and Wallis (2014) found no evidence for any representation of color in PFC in a similar task. Moreover, despite evidence for involvement of sensory areas in visual short-term memory processes (Pasternak & Greenlee, 2005; Harrison & Tong, 2009; Postle, 2016; Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017), no single neuron studies have investigated how maintenance of multiple stimuli in visual short-term memory affects representations in visual regions.

Here, we compare the neural correlates of visual short-term memory for multiple visual stimuli (colored squares) in LIP, PFC, and anterior inferotemporal cortex (IT) during a change localization task. In this task, we presented two displays of two to four colored squares separated by a brief delay, and the animal was rewarded for choosing the square that changed color. The three recorded regions are interconnected (Blatt, Andersen, & Stoner, 1990; Webster, Bachevalier, & Ungerleider, 1994; Saleem, Miller, & Price, 2014), and all are widely implicated in visual short-term memory processes (Fuster & Alexander, 1971; Fuster & Jervey, 1982; Passingham, 1975; Miyashita & Chang, 1988; Kowalska et al., 1991; Sawaguchi & Goldman-Rakic, 1991; Li et al., 1999; Petrides, 2000; Linden et al., 2003; Todd & Marois, 2004, 2005; Vogel & Machizawa, 2004; Voytek & Knight, 2010). Additionally,

anterior IT is critically involved in color perception (Heywood, Gaffan, & Cowey, 1995; Huxlin, Saunders, Marchionini, Pham, & Merigan, 2000). We investigated how stimulus load affected neural activity at stimulus encoding and over the delay period, and how these regions participated in determining which square changed color.

3.3 Methods

3.3.1 Behavioral Task

One adult female rhesus macaque (Macaca mulatta) was trained to perform a change localization task. All procedures followed the guidelines of the Massachusetts Institute of Technology Committee on Animal Care and the National Institutes of Health. The structure of the behavioral task is shown in Figure 3-1. After an initial fixation period of 1 second, the animal saw an array colored squares of 0.75 degrees of visual angle (DVA) at 5° of eccentricity. These squares then disappeared for a delay of 1.2 to 1.5 seconds before reappearing with a change to the color of one square. The animal received a juice reward for saccading to the changed square. If the animal broke fixation before the presentation of the test squares, the trial was aborted.

In all trials, a single square was shown in the hemifield ipsilateral to the recording sites, along the horizontal meridian. In the contralateral hemifield, squares could appear in four locations: along the horizontal meridian; in the upper and lower quarter field separated at an angle of 65° with respect to the horizontal meridian; in a "crowding" location in either the upper or lower quarter field at an angle of 40°, corresponding to a displacement of 2.2 DVA between their centers. The quarter field in which the crowding square was presented was chosen randomly by session. Data from trials on which the crowding square was presented are included only in Figure 3-10 and the corresponding results section. All other analyses were performed excluding these trials.

Each square had two possible associated colors. Colors were randomized across recording sessions. The minimum difference between the sample and test color of the changed square

was CIEDE2000 $\Delta E_{00} = 23.6$ (mean $\Delta E_{00} = 58.9$; max $\Delta E_{00} = 95.5$). Thus, color changes were very perceptible to a human observer (Yang et al., 2012). All colors were unique except for those of the square along the horizontal meridian and the crowding square, which shared the same color pair but were never presented simultaneously. One to three squares were shown in the contralateral hemifield, corresponding to a total load of 2 to 4. At test, the changed square was selected randomly. There were 88 possible sample stimulus displays (16 2-item displays, 40 3-item displays, and 32 4-item displays), of which 52 stimulus displays (12 2-item displays, 12 3-item displays, and 16 4-item displays) did not include the crowding square. The probability of each stimulus condition was uniform. The animal had to complete trials where each of the 88 displays was presented as a sample (not necessarily correctly) before any were shown again. The animal performed an average of 927 trials over the course of a session (inter-quartile range 864 – 1026) corresponding to 10.5 presentations of each sample display. In most analyses, we pooled across possible ipsilateral stimulus displays, yielding an average of 21 presentations per sample display.

3.3.2 Electrophysiological Recording

Recording sites were targeted using chambers that were custom-fit to the animal's skull based on a pre-implantation MRI (Mulliken et al., 2015). In IT, we localized a region rich in color-selective cells by coregistering the animal's anatomical MRI with functional MRIs obtained from other animals, graciously provided by Bevil Conway and Doris Tsao. We explored recording sites in preliminary recording sessions, characterized their properties during passive fixation, and selected the most color-selective locations for recordings during the task. Our recording sites most likely correspond to color patch ALc (Lafer-Sousa & Conway, 2013), on the lateral bank of the anterior middle temporal sulcus. Consistent with imaging results, highly color-selective sites were located medial to highly face-selective sites.

Epoxy-coated dura-piercing tungsten electrodes (FHC) were lowered into PFC and LIP using a custom-built grid and microdrive assembly that moved electrodes in pairs using a single screw. In inferotemporal cortex, a 32-contact V-Probe (Plexon) with 50 m spacing

was lowered using a Narishige MO-97A hydraulic microdrive. Recordings were performed using a Blackrock Microsystems Cerebus system. In sessions including IT recordings, signals were referenced to the shank of the V-Probe. In other sessions, signals were referenced to saline in the prefrontal chamber. In all sessions, the reference was tied to ground. Raw signals were digitized at 30 KHz and saved to disk for offline analysis.

Prior to spike detection, signals were denoised using independent components analysis (ICA) (Pedregosa et al., 2011; Hyvarinen, 1999). This process greatly reduced EMG and juicer-related artifacts, to the point that they were no longer audible in the processed signals. The denoising process consisted of four steps. First, we extracted a representative subset of the data, consisting of 3.5 second samples of 25 evenly spaced trials were extracted from the full recording session. Performing ICA on the entire session would be intractable, but since this subset is large (consisting of 2,625,000 samples), ICA yields nearly identical results. Second, we filtered this sample between 300 and 10,000 Hz using a second-order Bessel filter. Third, we whitened the data by PCA. Fourth, we applied ICA to the whitened signal. The initial mixing matrix was set to an adjusted transposed whitening matrix, such that, at the first step of the algorithm, the first $n - 1$ components were the signals recorded from each channel and the last component was the first principal component. This initialization process is not strictly necessary and similar results were obtained without it, but has the advantage that, after the algorithm converges, adjacent components generally reflect adjacent channels and the last component generally reflects the greatest contribution to the noise. Finally, we manually identified 1 to 3 noise components and removed them from the original signals. We separately denoised the signals recorded from tungsten electrodes and those from the V-Probe.

For the single-electrode recordings in PFC and LIP, we computed multi-unit activity by counting the number of negative deflections with an amplitude of at least 4×0.675 the median absolute signal in the data. For V-Probe recordings in IT, since spikes corresponding to individual neurons were generally present across 4 or more channels, we sorted units using Kilosort (Pachitariu, Steinmetz, Kadir, Carandini, & Harris, 2016) and manually inspected

the resulting clusters and their crosscorrelograms, discarding those that did not correspond to well-isolated units. We excluded units with a firing rate at sample presentation of less than 0.5 spikes per second, and units that were not isolated for at least 440 trials (5 trials of each stimulus condition). After these exclusions, we recorded 68 channels from LIP, 367 channels from PFC, and 155 single units from IT. For conciseness, we refer to both multi-unit activity and single-unit activity as "units" in the text below.

Since activity often drifted over the course of a session, we used a Gaussian process to rescale unit activity. We determined the spike count over the first second of the baseline period, and fit a Gaussian process with a radial basis function kernel to each unit using GPy (https://sheffieldml.github.io/GPy/), with a prior standard deviation of 1 and a prior length scale of 1 hour. For single-unit activity from IT, we used a Poisson likelihood and fit the model by Laplace approximation. Because multi-unit activity from PFC and LIP was generally overdispersed relative to a Poisson distribution but of sufficiently high rate that a normal approximation was sufficient, we used a normal likelihood. After fitting the Gaussian process model, we obtained smoothed estimates of the baseline rate on each trial, and multiplied the spiking activity by the ratio between the mean baseline activity over all trials and the smoothed activity for the given trial. Without rescaling, most analyses yielded similar results. However, in IT, firing rates during the baseline and delay periods were predictive of behavioral performance, even though there was no clear change in performance over time averaged over sessions. After smoothing, activity during these epochs bore no relationship to performance. While these effects could be physiological in origin, reflecting slow changes in brain state, they could also result from correlation between electrode drift due to animal motion and subsequent performance. Smoothing the data based on the baseline activity mitigates this potential confound.

3.3.3 Analysis of Proportion of Variance Explained (ε^2)

ε^2 is a debiased measure of the proportion of variance explained in an ANOVA. It is similar to, but less biased than, the more common η^2 and ω^2 measures, and is equivalent to adjusted

R^2 for the corresponding general linear model (Okada, 2013). It estimates:

$$\frac{\sigma^2_{\text{explained}}}{\sigma^2_{\text{total}}}$$

by replacing the numerator and denominator with unbiased sample estimators. For comparison of two models, the estimator of $\sigma^2_{\text{explained}}$ is:

$$\begin{aligned}\hat{\sigma}^2_{\text{explained}} &= MS_{\text{null}} - MS_{\text{full}} \\ MS_{\text{null}} &= \frac{RSS_{\text{null}}}{n - p_{\text{null}}} \\ MS_{\text{full}} &= \frac{RSS_{\text{full}}}{n - p_{\text{full}}}\end{aligned}$$

MS_{null} and MS_{full} are unbiased estimates of the variance of the residuals of the null and full models, respectively; RSS_{null} and RSS_{full} are their respective sums of squares; p_{null} and p_{full} are their parameter count including intercept; and n is the total number of observations. Since all sample stimulus conditions were presented a nearly equal number of times, ANOVAs involving them are nearly orthogonal, and $MS_{\text{null}} - MS_{\text{full}}$ is nearly equal to the mean square of the model including factors in MS_{full} that are not included in MS_{null}.

ε^2 estimates the proportion of the total trial-to-trial variance that the given model can explain, with σ^2_{total} given by the unbiased sample variance of the dependent variable $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}$. Partial ε^2 (ε^2_p) estimates the proportion of the residual variance of the null model that can be explained by the full model, with σ^2_{total} estimated by MS_{null}. When comparing variance explained by different factors within the same data, we computed ε^2, so that the denominator was the same between factors. When comparing variance explained by the same factors in different data, we computed ε^2_p. When we sought to compare ε^2_p across factors, but to exclude contributions to the variance of factors not compared, we estimated σ^2_{total} using the mean square of a different, third model. Table 3.1 shows how we computed ε^2 for all analyses.

ε^2 has significant advantages compared to alternative strategies for measuring information

in spiking activities. It allows testing of complex hypotheses with minimal bias, and since it is computed on individual units, it is easy to construct error bars and confidence intervals for the population. Fitting ordinary linear models produces parameter estimates with higher mean squared error than could be achieved with regularization or with more complex models. However, the predictive power of these models is of little interest. As discussed in Shmueli (2010), explanation and prediction are different goals that often demand different models. For our purposes, it is of great utility that ε^2 is a nearly unbiased estimator.

When performing permutation tests of ε^2, we used the same models and shuffled according to the null model, with the exception of analyses performed over the test period. In these cases, no exact permutation test can be constructed, since S_n, S_n^{test}, and $S_n : S_n^{\text{test}}$ are all dependent and their dependency structure does not permit permutation (see Table 3.1 for nomenclature). To test for an effect of sample color, we computed ε_p^2 by comparing the variance of the model $S_2 \times S_3 \times S_5 + S_2^{\text{test}} \times S_3^{\text{test}} \times S_5^{\text{test}}$ to $S_2^{\text{test}} \times S_3^{\text{test}} \times S_5^{\text{test}}$, and permuted within test displays. To test for an effect of test color, we compared the same full model to $S_2 \times S_3 \times S_5$, and permuted within sample displays. To test for an effect of the changed square, we computed ε_p^2 for a model $P_2 \times P_3 \times P_5 + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$ relative to a model including $P_2 \times P_3 \times P_5$. Since the probability of a given square change was uniform for a given position condition, we permuted within position conditions. Values of ε_p^2 plotted in Figure 3-11B correspond to this model comparison.

When computing the linear change in ε_p^2 with load, shown in Figure 3-8, we fit models for each load condition as in Figure 3-7. For each unit, we fit the ordinary least squares model:

$$y_l = \beta_0 + (l-2)\beta_1 + \varepsilon_l$$

where l is the total load, i.e., 2, 3, or 4; y_l is the value of ε_p^2 at load l; and ε_l is the error term. Since OLS produces unbiased parameter estimates and values of y_l are nearly unbiased, β_1 is a nearly unbiased estimator of the increase in variance explainable by square color with each square added to the display. Its magnitude depends on the overall degree of color selectivity, but we sought solely to compare its sign between regions. As shown in Figure 3-8, β_1 was

Item	Name	Trials	Full Model	Null Model	Denom.
Table 3.2	Stimulus Condition	No S_4	$S_1 \times S_2 \times S_3 \times S_5$	**1**	**1**
Table 3.2	Contralateral Color	No S_4	$S_1 + S_2 \times S_3 \times S_5$	$S_1 + P_2 \times P_3 \times P_5$	**1**
Table 3.2	Ipsilateral Color	No S_4	$S_1 + S_2 \times S_3 \times S_5$	$S_2 \times S_3 \times S_5$	**1**
Fig. 3-3	Stimulus Condition	No S_4	$S_2 \times S_3 \times S_5$	**1**	**1**
Fig. 3-3	Color	No S_4	$S_2 \times S_3 \times S_5$	$P_2 \times P_3 \times P_5$	**1**
Fig. 3-4	Position	Load 2, no S_4	$S_2 + S_3 + S_5$[a]	$S_2 + S_3 + S_5$[b]	**1**
Fig. 3-4	Color	Load 2, no S_4	$S_2 + S_3 + S_5$[a]	$P_2 + P_3 + P_5$	**1**
Fig. 3-5	Position	Passive fixation	Position × Color[c]	Color + Position : Color	**1**
Fig. 3-5	Color	Passive fixation	Position × Color[c]	Position+Position : Color	**1**
Fig. 3-5	Interaction	Passive fixation	Position × Color[c]	Position + Color	**1**
Fig. 3-7	Load 2	Load 2, no S_4	$S_2 + S_3 + S_5$[a]	$P_2 + P_3 + P_5$	$P_2 + P_3 + P_5$
Fig. 3-7	Load 3	Load 3, no S_4	$S_2 \times S_3 \times S_5$	$P_2 \times P_3 \times P_5$	$P_2 \times P_3 \times P_5$
Fig. 3-7	Load 4	Load 4, no S_4	$S_2 \times S_3 \times S_5$	$P_2 \times P_3 \times P_5$	$P_2 \times P_3 \times P_5$
Fig. 3-10B	Load 3 Uncrowded	Only S_3 & S_5	$S_3 \times S_5$	**1**	**1**
Fig. 3-10B	Load 3 Crowded	Only S_4 & S_5	$S_4 \times S_5$	**1**	**1**
Fig. 3-10B	Load 4 Uncrowded	S_2, S_3, & S_5	$S_2 + S_3 \times S_5$	S_2	S_2
Fig. 3-10B	Load 4 Crowded	S_2, S_4, & S_5	$S_2 + S_4 \times S_5$	S_2	S_2
Fig. 3-10C	Linear Uncrowded	S_3 & S_5	$S_2 + P_2 \times S_3 \times S_5$	$S_2 + P_2 \times (S_3 : S_5)$	S_2
Fig. 3-10C	Linear Crowded	S_4 & S_5	$S_2 + P_2 \times S_4 \times S_5$	$S_2 + P_2 \times (S_4 : S_5)$	S_2
Fig. 3-10C	Nonlinear Uncrowded	S_3 & S_5	$S_2 + P_2 \times S_3 \times S_5$	$S_2 + P_2 \times (S_3 + S_5)$	S_2
Fig. 3-10C	Nonlinear Crowded	S_4 & S_5	$S_2 + P_2 \times S_4 \times S_5$	$S_2 + P_2 \times (S_4 + S_5)$	S_2
Fig. 3-11A	Test Color	No S_4	$S_2 \times S_3 \times S_5 + S_2^{\text{test}} \times S_3^{\text{test}} \times S_5^{\text{test}} + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$S_2 \times S_3 \times S_5 + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$P_2 \times P_3 \times P_5$
Fig. 3-11A	Sample Color	No S_4	$S_2 \times S_3 \times S_5 + S_2^{\text{test}} \times S_3^{\text{test}} \times S_5^{\text{test}} + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$S_2^{\text{test}} \times S_3^{\text{test}} \times S_5^{\text{test}} + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$P_2 \times P_3 \times P_5$
Fig. 3-11A	Changed Square	No S_4	$S_2 \times S_3 \times S_5 + S_2^{\text{test}} \times S_3^{\text{test}} \times S_5^{\text{test}} + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$S_2 \times S_3 \times S_5 + S_2^{\text{test}} \times S_3^{\text{test}} \times S_5^{\text{test}}$	$P_2 \times P_3 \times P_5$
Fig. 3-12	Changed Square	Load 3 or 4, no S_4, contralateral target, correct or incorrect	$P_2 \times P_3 \times P_5 + S_2 + S_3 + S_5 + S_2^{\text{test}} + S_3^{\text{test}} + S_5^{\text{test}} + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$P_2 \times P_3 \times P_5 + S_2 + S_3 + S_5 + S_2^{\text{test}} + S_3^{\text{test}} + S_5^{\text{test}}$	$P_2 \times P_3 \times P_5$
Fig. 3-12	Sample Color	Load 3 or 4, no S_4, contralateral target, correct or incorrect	$P_2 \times P_3 \times P_5 + S_2 + S_3 + S_5 + S_2^{\text{test}} + S_3^{\text{test}} + S_5^{\text{test}} + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$P_2 \times P_3 \times P_5 + S_2^{\text{test}} + S_3^{\text{test}} + S_5^{\text{test}} + S_2 : S_2^{\text{test}} + S_3 : S_3^{\text{test}} + S_5 : S_5^{\text{test}}$	$P_2 \times P_3 \times P_5$

Table 3.1: ANOVA Factors. Factors in ANOVAs used to compute ε^2. **Figure** indicates the figure in which the ANOVA was presented. **Model Name** indicates the name of the ANOVA in the corresponding figure. **Trials** indicates the trials to which the ANOVA was fit. **Full Model** is the model from which MS_{full} was computed. The **Null Model** is the model from which MS_{null} was computed, which was always nested inside the full model. **Denom.** is the model used to compute σ^2_{total} above. **1** denotes the intercept-only model. "+" denotes a linear effect; ":" denotes an interaction; "×" denotes a linear effect and its interaction, i.e. $a \times b \equiv a + b + a : b$. S_1 is the ipsilateral square, which was always located along the horizontal meridian. Remaining squares were all contralateral, and are numbered according to their absolute angular displacement from the uncrowded quarter field. S_2 was in the uncrowded quarter field; S_3 was along the horizontal meridian; S_4 was the "crowding" square at 40° from the horizontal meridian; and S_5 was at 65° from the horizontal meridian that is adjacent to S_4. S_n is a three-level factor, reflecting whether a square was not present, color A, and color B. P_n is a two-level factor that reflects whether square S_n was presented. S_n^{test} is the same as S_n, but for squares presented at test.

[a] In the load 2 condition, only one contralateral square was presented, so there are no interaction terms.

[b] Responses to the two possible colors of S_n were constrained to sum to zero.

[c] During passive fixation, we presented only a single square, but manipulated its position and color factorially.

significantly greater than zero in PFC during the sample and delay, but significantly less than zero in IT, and there was a significant interaction.

When determining the latency of modulation by the changed square (Figure 3-11B), we convolved firing rates with a raised cosine function of 20 ms full width at half maximum. Unlike a Gaussian, this raised cosine function has finite support on an interval of 60 ms. Thus, values of ε_p^2 reflect activity at most 30 ms after a given time point. We computed values of ε_p^2 for 1000 permutations within positions over time points from 0 to 161 ms after stimulus onset, and fixed the significance threshold as the 95% quantile of the maxima over time of the permutation. We determined the latency as the first time point at which the observed value of ε_p^2 exceeded this threshold. We tested for a significant difference difference between regions by performing the same procedure for 1000 realizations of the data, resampling with replacement among units, and computing a two-sided bootstrap p-value:

$$p = \frac{2}{1000}\min\left(\sum_{i=1}^{1000}\{a_i \geq b_i\}, \sum_{i=1}^{1000}\{a_i \leq b_i\}\right)$$

where a_i is replicate i from region a and b_i is replicate i from region b.

When testing for differences between correct and incorrect trials (Figure 3-12), we restricted analysis to trials where the load was 3 or 4, since the animal rarely responded incorrectly on load 2 trials. We additionally required that the target square was contralateral, since the animal's performance was higher for ipsilateral targets at all load conditions (Figure 3-2). Because these restrictions left few incorrect trials of any given type, we fit a simplified model (Table 3.1). We constructed permutations by shuffling the correct and incorrect trial labels within load conditions, such that the numbers of correct and incorrect load 3 and load 4 trials for each permutation were equal to the numbers of correct and incorrect load 3 and load 4 trials for the true data. Although ε_p^2 was never significantly higher on incorrect trials than correct trials, we computed 2-sided p-values by first computing the maximum and minimum of the difference in ε_p^2 over time points from -100 to 300 ms after test onset for each permutation, then comparing the true values at each time point against these permutation minima and maxima and adjusting for the two tests performed.

3.3.4 Classification Analysis

In order to confirm that the load effects that we observe do not result solely from our choice of ε^2 as a measure of color selectivity, we tested these effects using a different analysis strategy. We trained linear support vector machines (SVMs) to distinguish between two stimulus displays where squares were located in the same positions and only one contralateral square differed in color, for all valid pairs of uncrowded stimulus displays. We first restricted our analyses to units for which we recorded at least 14 trials of each contralateral stimulus condition. Next, we split the data into 14 possible train/test splits, each consisting of 13 training trials and 1 test trial for each contralateral stimulus condition. For each split, we trained SVMs for each pair of square colors as follows. First, we computed the mean firing rate for each unit over training trials, and the pooled standard deviation for each unit as the square root of the mean within-condition variance of the training trials. We z-scored the training and test sets using these estimates. Next, we estimated the optimal SVM soft margin parameter C based on 3-fold cross-validation within the training set. Finally, we trained an SVM on the entire training set using the value of C that produced the best results on the cross-validation sets, and tested it on the test set. We computed the classifier performance as the proportion of correct classifications over all test sets and square color pairs. We trained a classifier for each unit individually, as well as for 500 random subsets of 2, 3, 4, 6, 8, 11, 16, 23, 32, 45, 64, 91, and 128 units. In PFC, we also investigated subsets of 181 and 256 units.

Computing accurate confidence intervals for classification analysis is non-trivial. While confidence intervals could in principle be computed as the proportion of correct responses over train/test splits, it is unclear whether such confidence intervals would be valid, since the training data are not independent between splits. Additionally, when comparing results between regions, confidence intervals should remain valid for different units recorded from these regions, not simply for different trials recorded from the same units. Bootstrapping over units is also problematic, since redundant features are less informative than unique features sampled from the same distribution. Indeed, we found that bootstrapping over all

units yielded classifier performance that was substantially lower than the true performance. We thus computed classifier performance over subsamples of units drawn without replacement, and estimated the variance with a correction for finite population (Politis, Romano, & Wolf, 1999). For a vector $\theta_1, ..., \theta_m$ of subsamples of k units from a sample of n units, we approximated the variance of $\bar{\theta}$ as:

$$\hat{\sigma}^2 = \frac{k}{m(n-k)} \sum_{i=1}^{m} (\theta_i - \bar{\theta})^2$$

This formula is equivalent to the equivalent to that of the delete-d jackknife variance estimator (Shao & Wu, 1989). In general, subsampling is applicable to a wider variety of circumstances than bootstrapping (Andrews, 1999; Politis et al., 1999). This estimator generally produces larger variance estimates than a bootstrap estimator, especially as $k \to n$, although the two methods are asymptotically equivalent as $k/n \to 0$. Although we approximated the distribution of the mean over subsamples with a normal distribution, results closely resemble those obtained by estimating the distribution based on empirical quantiles of $\theta_i - \bar{\theta}$, suggesting that a normal approximation is adequate.

We fit the change in classification accuracy with load using a logistic regression model:

$$y_l \sim \text{Binomial}\left(n_l, \frac{1}{1 + e^{-(\beta_0 - (l-2)\beta_1)}}\right)$$

where y_l is the number of held out trials that were correctly classified at load l, and n_l is the total number of held out trials. β_0 estimates the log odds ratio of correct classification at load 2. β_1 estimates the change in log odds ratio per square added to the display. If β_1 is less than 0, classification performance decreases with load; if β_1 is greater than zero, classification performance increases with load; and if β_1 is 0, classification performance has no linear relationship with load. We fit this model for every subsample of units, took the mean of β_1 across subsamples, and estimated its variance $\hat{\sigma}^2_{\beta_1}$ using the formula above, and performed hypothesis tests by approximating the distribution of β_1 by its mean across subsamples $\hat{\beta}_1$ and its variance $\hat{\sigma}^2_{\beta_1}$. Specifically, when testing whether β_1 was significantly

different from zero, we computed the p-value as:

$$p = 2\left(1 - \Phi\left(\frac{\left|\bar{\beta}_1\right|}{\hat{\sigma}_{\beta_1}}\right)\right)$$

where Φ is the cumulative distribution function of the standard normal distribution. When testing whether β_1 differed between sets of subsamples a and b, we computed the p-value as:

$$p = 2\left(1 - \Phi\left(\frac{\left|\bar{\beta}_1^a - \bar{\beta}_1^b\right|}{\sqrt{\hat{\sigma}^2_{\beta_1^a} + \hat{\sigma}^2_{\beta_1^b}}}\right)\right)$$

3.4 Results

The animal performed a change localization task, shown in Figure 3-1A. Animals were required to encode and maintain an array of colored squares in visual short-term memory. We manipulated the number of contralateral squares as well as their locations. Two, three, or four squares were presented simultaneously. In the hemifield ipsilateral to the recording site, one square was present along the horizontal meridian on every trial. In the hemifield contralateral to the recording site, and one, two, or three squares were presented. The seven position conditions are shown in Figure 3-1B. Since we also varied the colors of the squares, there were a total of 52 sample stimulus conditions. On average, the animal performed 10.5 trials for each condition. As the animal performed the task, we recorded from the lateral intraparietal area (LIP), prefrontal cortex (PFC), and a region in anterior inferotemporal cortex (IT) chosen for its color selectivity.

3.4.1 Behavioral Performance

Figure 3-2 shows the animal's performance. As expected, performance decreased with load (all $p < 10^{-52}$, Fisher's exact test; Figure 3-2A). We measured the animal's ability to detect the target square as conditional mutual information (CMI) between the target square location and chosen square location given the shown square locations. CMI increased from load 2 to 3

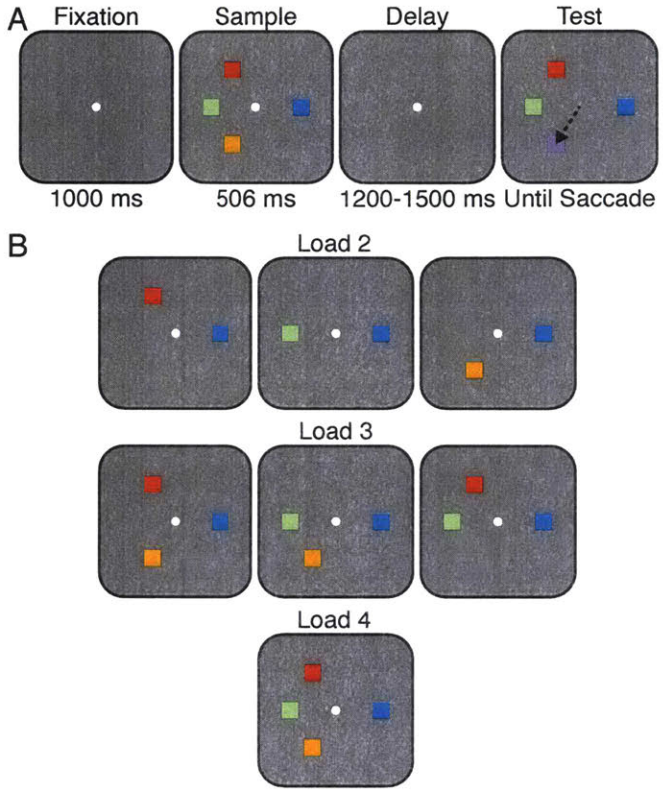


Figure 3-1: Change Localization Task. **A**: After fixating for 1000 ms, the animal saw an array of colored squares at 5 degrees of eccentricity. These squares then disappeared for a delay of 1.2 to 1.5 seconds before reappearing with a change to the color of a random square. The animal was rewarded for saccading to the changed square. **B**: A single colored square was presented in the same position within the hemifield ipsilateral to the recording sites on every trial. The number of contralateral squares was varied from one to three. Each square could be one of two colors. Possible colors were different for every square, and were randomized between sessions.

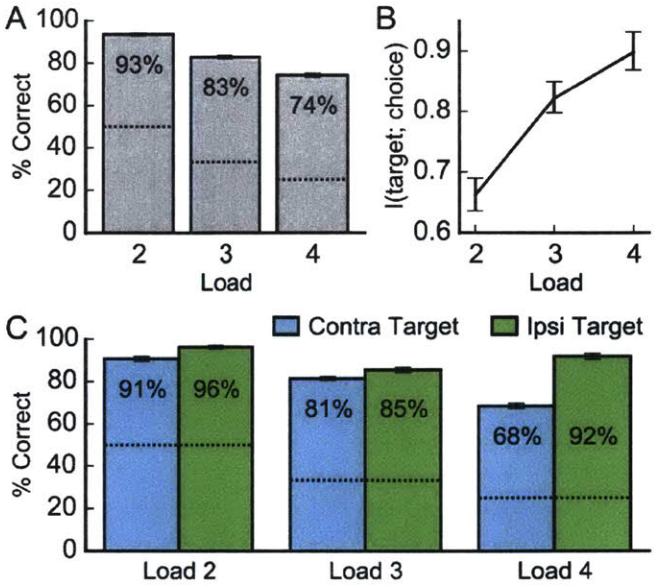


Figure 3-2: Behavioral Performance. **A**: Percentage of correct trials by load condition. Dotted lines indicate chance performance. **B**: Mutual information between the location of the changed (target) square and the animal's choice. Information increased with load. **C**: Percentage of correct trials by load condition and target side. "Contra Target" indicates that the changed square was in the hemifield contralateral to the recording site, in which the number of squares varied from 1 to 3 at loads 2 to 4. "Ipsi Target" indicates that the changed square was in the hemifield ipsilateral to the recording site, which always contained a single square. Dotted lines indicate chance performance.

and from load 3 to 4 (both $p < 10^{-5}$, bootstrap test; Figure 3-2B). The increase in CMI was larger from load 2 to 3 than from load 3 to 4 ($p = 0.0006$, bootstrap test), providing possible evidence of the plateau observed by Buschman et al. (2011). As in that study, performance was higher when the ipsilateral square changed than when a contralateral square changed at all loads (all $p < 10^{-8}$, Fisher's exact test; Figure 3-2C).

3.4.2 Selectivity for Position and Color

Activity in all three recorded regions differed by the stimulus. To measure selectivity, we computed ε^2, a debiased measure of the proportion of variance in the firing rate explained, and compared the value against a permutation distribution. During the sample period, from 80 to 505 ms after sample onset, activity in all three regions differed depending on

Selectivity for	LIP ($n = 68$) Units	$\bar{\varepsilon}^2$	p	PFC ($n = 367$) Units	$\bar{\varepsilon}^2$	p	IT ($n = 155$) Units	$\bar{\varepsilon}^2$	p
Sample									
Stimulus Condition	21	0.073	$< 10^{-4}$	116	0.074	$< 10^{-4}$	95	0.152	$< 10^{-4}$
Contralateral Color	5	0.0059	0.0001	45	0.012	$< 10^{-4}$	88	0.105	$< 10^{-4}$
Ipsilateral Color	1	0.0001	0.67	4	0.0004	0.008	6	0.002	0.005
Delay									
Stimulus Condition	9	0.022	$< 10^{-4}$	103	0.054	$< 10^{-4}$	4	0.0054	0.004
Contralateral Color	0	0.0003	0.58	27	0.0065	$< 10^{-4}$	5	0.0033	0.0002
Ipsilateral Color	0	-0.0002	0.67	5	0.0002	0.11	3	0.0001	0.32

Table 3.2: Selectivity for Stimulus Condition and Color. Number of units in each region selective for stimulus condition, contralateral color, and ipsilateral color at $\alpha = 0.01$; mean values of ε^2; and p-values for mean ε^2. Selectivity was assessed by computing ε^2 for the given attribute and comparing it against a permutation distribution (see Methods).

which of the 52 possible stimulus displays had been presented (Table 3.2). Additionally, within position conditions, the colors of the contralateral squares modulated activity in all regions. The color of the ipsilateral square also modulated activity in PFC and IT, but this modulation was weak. Significantly fewer units were modulated by ipsilateral color than contralateral colors (PFC: 1% vs. 12%, $p < 10^{-9}$; IT: 4% vs. 57%, $p < 10^{-25}$, Fisher's exact test). Ipsilateral color also accounted for little variance in firing rate relative to contralateral color (PFC: $\bar{\varepsilon}^2 = 0.0004$ vs. 0.013, $t(366) = 7.1$, $p < 10^{-11}$; IT: $\bar{\varepsilon}^2 = 0.002$ vs. 0.106, $t(154) = 10.0$, $p < 10^{-17}$, t-test).

All three regions exhibited stimulus-selective delay period activity in the window from 300 to 1200 ms after stimulus offset (Table 3.2). However, only PFC and IT were significantly modulated by the contralateral square color. Modulation by contralateral color was weak in both PFC and IT, with no significant difference between the regions (PFC $\bar{\varepsilon}^2 = 0.0065$, IT $\bar{\varepsilon}^2 = 0.0033$; $p = 0.11$, unequal variance t-test), even though modulation by contralateral color at sample presentation was substantially stronger in IT than PFC ($\bar{\varepsilon}^2 = 0.105$ vs. 0.012; $p < 10^{-15}$). Because modulation by ipsilateral color was weak during the sample and non-significant during the delay, we pooled over ipsilateral color conditions when performing all analyses below.

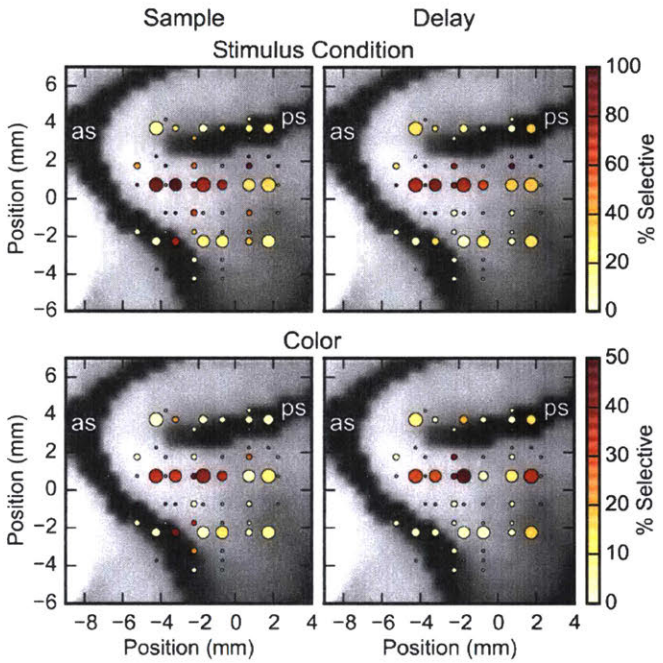


Figure 3-3: Topographical Distribution of Selectivity in PFC. Proportions of penetrations yielding any selectivity for stimulus condition (top) or selectivity for contralateral stimulus colors (bottom) during the sample (left) or delay (right), as assessed by comparing ε^2 against a permutation distribution at $\alpha = 0.01$. The size of circles indicates the number of penetrations at the given site. Anatomy was reconstructed by co-registering a post-implantation MRI to a pre-implantation MRI and computing a map of the distance to cortex along the axis of the recording chamber. as = arcuate sulcus, ps = principal sulcus.

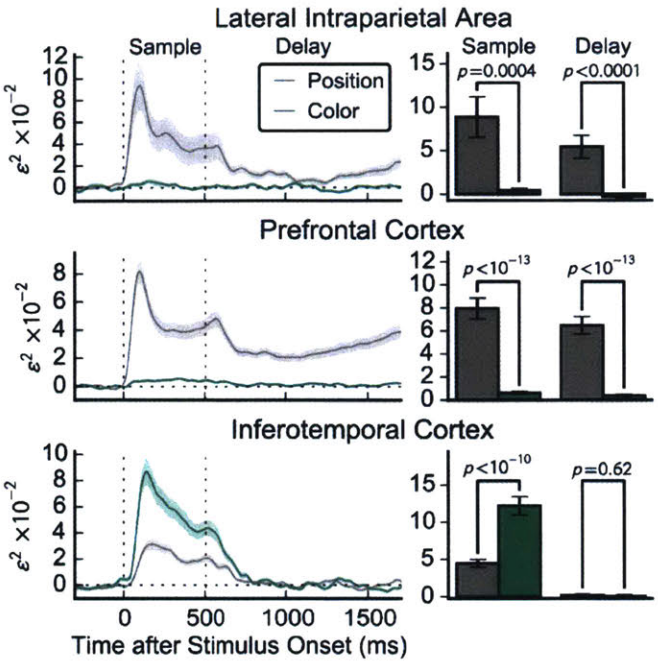


Figure 3-4: Selectivity for Position and Color. The debiased proportion of variance explained (ε^2) by position and color in each recorded region. Left: Proportion of variance explained over time. Dotted vertical lines indicate the onset and offset of the sample. Firing rates were smoothed with a Gaussian of 50 ms FWHM. Right: Proportion of variance explained over the sample (60 to 505 ms after stimulus onset) and delay (300 to 1200 ms after stimulus offset) epochs. p-values indicate significance by dependent samples t-test. Error bars are $\pm$ SEM.

Within PFC, stimulus selectivity differed greatly across recording sites. Figure 3-3 shows the proportion of penetrations that yielded activity selective to stimulus condition and color for different sites recorded in PFC. Within individual sites, selectivity was similar between sample and delay, and between stimulus condition and color. A substantially larger proportion of electrode penetrations along the caudal portion of the arcuate yielded selective activity as compared to the rest of PFC. This region is near the junction of areas 8av, 46v, and 45a (Saleem et al., 2014), in a location consistent with the ventral prearcuate region identified by Bichot, Heard, DeGennaro, and Desimone (2015), and the region in which Lundqvist et al. (2016) found the strongest modulation of spiking activity and gamma modulation during a sequential multiple-item visual short-term memory task.

Using only two-item trials, we compared selectivity for the position of the contralateral

square with selectivity for its color. Figure 3-4 shows ε^2 across the three recorded regions. At sample presentation, units in all three regions were modulated by both the positions and the colors of the squares (all $p < 0.005$, permutation test of mean ε^2). LIP and PFC were more strongly modulated by the positions of the presented squares than their colors (LIP: $t(67) = 3.7$, $p = 0.00005$; PFC: $t(366) = 8.0$, $p < 10^{-13}$, t-test; Figure 3-4). By contrast, in inferotemporal cortex, neurons were more strongly modulated by color than position ($t(154) = 7.1$, $p < 10^{-10}$). During the delay, PFC remained more strongly modulated by position than color (PFC: $t(366) = 7.9$, $p < 10^{-13}$). Position selectivity persisted in LIP, and was significantly stronger than the insignificant color selectivity ($t(67) = 4.4$, $p < 0.0001$). During two-item delays, IT was not significantly selective for either position or color (both $p > 0.4$). Thus, PFC and LIP preferentially represented stimulus positions, whereas IT preferentially represented stimulus colors.

Since colors in the change localization task were different for every square, position selectivity could potentially reflect a unit's preference for colors shown at some squares over colors shown at other squares. To fully disentangle position from color selectivity, in a subset of sessions, we recorded from PFC and IT while the animal passively fixated single squares of all possible colors at all possible locations. Both regions were selective to position and color (all $p < 0.0001$, permutation test of ε^2; Figure 3-5). IT, but not PFC, also showed a significant interaction (IT: $p < 0.0001$, PFC: $p = 0.07$). As in the change localization task, more variance was explained by position than color or interactions in PFC (position vs. color: $t(135) = 5.6$, $p < 10^{-6}$; position vs. interaction: $t(135) = 6.0$, $p < 10^{-7}$, t-test), and more variance was explained by color than position or interactions in IT (color vs. position: $t(126) = 8.8$, $p < 10^{-14}$; color vs. interaction: $t(126) = 8.5$, $p < 10^{-13}$, t-test).

3.4.3 Modulation by Load

Firing rates at sample were significantly modulated by load in PFC, but no significant differences were present in LIP or IT. Figure 3-6 shows average unit activity over time in the three recorded regions. In PFC, the strength of the initial response transient (from 100

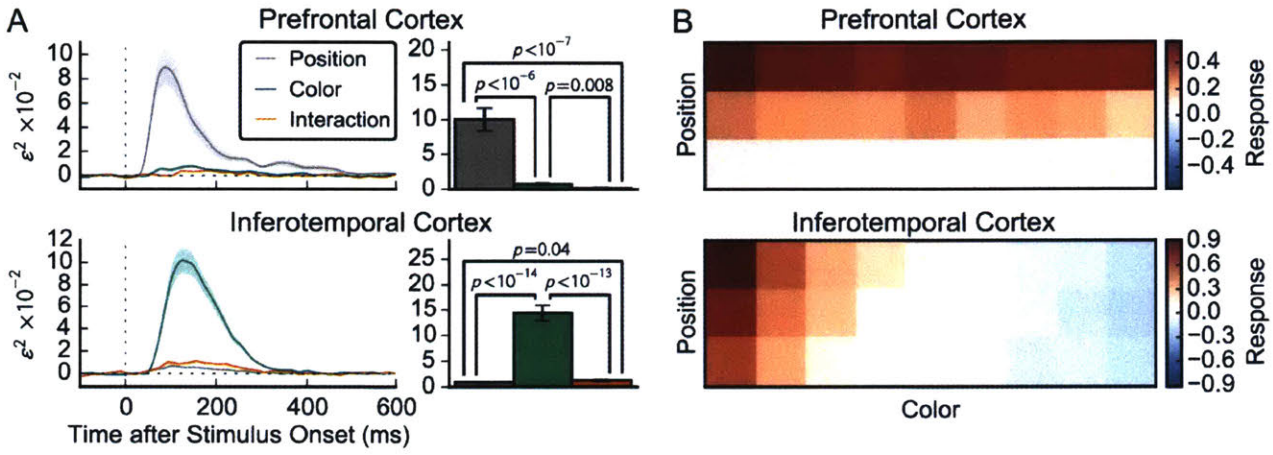


Figure 3-5: Selectivity for Position and Color During Passive Fixation. **A**: The debiased proportion of variance explained (ε^2) by position, color, and their interaction in PFC ($n = 136$ channels) and IT ($n = 127$ units). Left: Proportion of variance explained over time. Dotted vertical line indicates the onset of the stimulus. The stimulus was presented for 106 ms, followed by a 106 ms blank period. Firing rates were smoothed with a Gaussian of 30 ms FWHM. Right: Proportion of variance explained from 60 to 180 ms after stimulus onset. p-values indicate significance by dependent samples t-test. Error bars are ± SEM. **B**: Average responses to positions and colors, z-scored with respect to activity during the baseline (80 to 150 ms after stimulus offset). For each unit, prior to averaging, positions were sorted according to the maximum response over colors, and colors were sorted according to the maximum response over positions. To avoid circularity, sort order was determined on odd presentations, and responses were measured on even presentations.

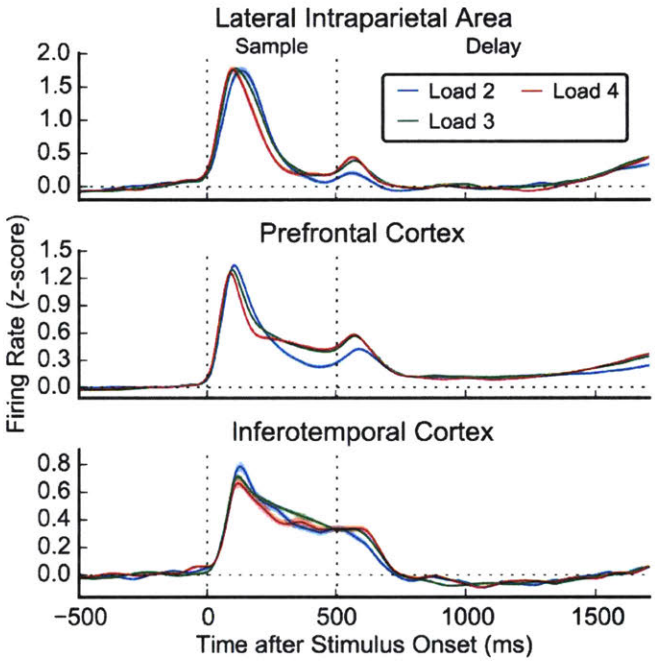


Figure 3-6: Changes in Firing Rate with Load. Firing rates in LIP, PFC, and IT, baseline-subtracted and normalized by the baseline standard deviation, and smoothed with a Gaussian of 75 ms FWHM. Error bars are ± SEM after removing firing rates averaged across loads by unit.

to 200 ms after sample onset) decreased with load ($F(2,732) = 36.6$, $\varepsilon = 0.59$, $p < 10^{-9}$, repeated measures ANOVA with Greenhouse-Geisser correction). During this period, firing rates were higher at 2 compared to loads 3 and 4 (load 2 vs. 3: $t(366) = 5.0$, $p < 10^{-6}$; load 2 vs. 4: $t(366) = 6.4$, $p < 10^{-9}$), and at load 3 compared to load 4 ($t(366) = 6.4$, $p < 10^{-9}$, t-test). However, the strength of the sustained response (from 300 to 505 ms after sample onset) increased with load ($F(2,732) = 16.4$, $\varepsilon = 0.59$, $p < 10^{-4}$). Firing rates were higher at loads 3 and 4 compared to load 2 (load 3 vs. 2: $t(366) = 5.0$, $p < 10^{-6}$; load 4 vs. 2: $t(366) = 3.9$, $p = 0.0001$). There was no significant difference between load 3 and 4 ($t(366) = 1.1$, $p = 0.24$).

After sample presentation, firing rate modulation by load disappeared, but was reestablished just before the onset of the test stimulus (Figure 3-6). In the middle of the delay (from 300 to 900 ms after sample offset), there was no difference in firing rate between load conditions in any region (all $p > 0.27$, repeated measures ANOVA with Greenhouse-Geisser correction). However, over the 200 ms before test stimulus presentation, firing rates in all regions increased (LIP: $t(67) = 5.3$, $p < 10^{-5}$; PFC: $t(366) = 7.1$, $p < 10^{-11}$; IT: $t(154) = 4.8$, $p < 10^{-5}$). During this period, LIP and PFC were modulated by load (LIP: $F(2,134) = 8.1$, $\varepsilon = 0.61$, $p = 0.003$; PFC: $F(2,732) = 17.1$, $\varepsilon = 0.62$, $p < 10^{-5}$). In PFC, firing rates were higher at loads 3 and 4 compared to load 2 (load 2 vs. 3: $t(366) = 4.0$, $p < 0.0001$; load 2 vs. 4: $t(366) = -4.4$, $p < 0.0001$) and at load 4 compared to load 3 ($t(366) = 3.3$, $p = 0.001$). In LIP, a firing rates were greater at load 3 than load 2 ($t(67) = 3.7$, $p = 0.0004$), and marginally greater at load 4 than load 2 ($t(67) = 2.6$, $p = 0.01$), but not significantly different between loads 3 and 4 ($t(67) = 0.39$, $p = 0.70$). There was no significant modulation by load in IT in any epoch. However, IT firing rates were suppressed relative to baseline in the middle of the delay ($t(154) = 3.6$, $p = 0.0005$), in contrast to PFC firing rates, which were enhanced ($t(366) = 3.7$, $p = 0.0003$). No significant difference between baseline and the middle of the delay was observed in LIP ($t(67) = 0.1$, $p = 0.90$). Thus, although IT activity during the delay bore little relationship to the contents of visual short-term memory, it was nonetheless modulated by the task.

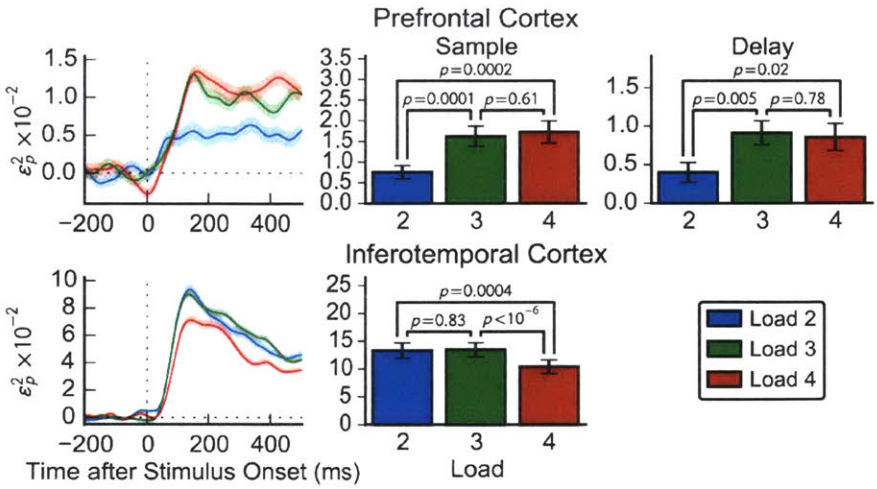


Figure 3-7: Changes in Color Selectivity with Load. The debiased proportion of variance explained (ε_p^2) by contralateral square colors in PFC and IT. Left: Proportion of variance explained over the sample period. Firing rates were smoothed with a Gaussian of 75 ms FWHM. Right: Proportion of variance explained averaged over the sample (80 to 505 ms after stimulus onset) and delay (300 to 1200 ms after stimulus offset) epochs. p-values indicate significance by dependent samples t-test. Error bars are $\pm$ SEM after removing color selectivity averaged across loads by unit.

Beyond simple investigation of firing rates, we also sought to determine how information about the colors of the presented stimuli varied with load. Since we recorded few LIP channels selective for color at sample or delay, and few IT units selective for color during the delay, we restricted this analysis to the sample and delay periods in PFC and the sample period in IT. Figure 3-7 shows the proportion of variance explained by the colors of all stimuli (ε_p^2), including possible interactions between stimuli, in PFC and IT. At sample presentation and during the delay, color selectivity was greater in PFC at loads 3 and 4 than at load 2 (load 3 vs. 2 at sample: $t(366) = 3.8$, $p = 0.0001$; load 4 vs. 2 at sample: $t(366) = 3.8$, $p = 0.0002$; load 3 vs. 2 during delay: $t(366) = 2.8$, $p = 0.005$; load 4 vs. 2 during delay: $t(366) = 2.3$, $p = 0.02$, t-test), but there was no significant difference between loads 4 and 3 (sample: $t(366) = 0.5$, $p = 0.61$, delay: $t(366) = -0.3$, $p = 0.78$). In IT, color selectivity did not increase with load. It was not significantly different between loads 2 and 3 ($t(154) = -0.2$, $p = 0.83$), but decreased at load 4 (load 2 vs. 4: $t(154) = 3.6$, $p = 0.0004$; load 3 vs. 4: $t(154) = 5.3$, $p < 10^{-6}$).

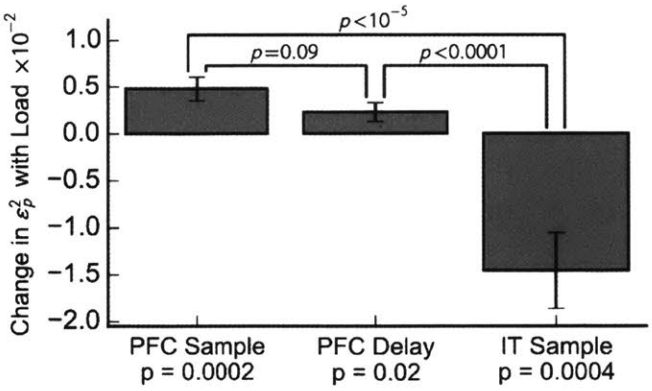


Figure 3-8: Scaling of Color Selectivity with Load. Change in debiased proportion of variance explained (ε_p^2) by contralateral square colors in PFC and IT with number of squares. p-value for PFC sample versus delay reflects significance by a dependent samples t-test. p-values for PFC versus IT reflect significance by an unequal variance t-test. p-values at bottom reflect significant differences from zero according to a one-sample t-test. Error bars are ± SEM.

Thus, while PFC carried more color information as additional squares were added to the display, IT carried less. To verify that this regional difference was statistically significant, we computed the change in the proportion of variance explained (ε_p^2) by color with load in PFC and IT for each unit by linear regression, then averaged these values across the population of units (Figure 3-8). The proportion of variance explained increased with load in PFC during the sample ($t(366) = 3.8$, $p = 0.0002$, t-test) and delay ($t(366) = 2.3$, $p = 0.02$), but there was no significant difference between effects during the two epochs ($t(366) = 2.0$, $p = 0.09$, t-test). In IT, it instead decreased with load during the sample ($t(154) = 3.6$, $p = 0.0004$). Changes with load were significantly different between regions (PFC sample vs. IT sample: $t(185.6) = 4.6$, $p < 10^{-5}$; PFC delay vs. IT sample: $t(173.8) = 4.0$, $p < 0.0001$, unequal variance t-test). We observed a similar pattern if only units selective in the given epoch were included in this analysis.

We also confirmed this difference between load effects in PFC and IT in a different way, by performing a classification analysis. Holding the colors of other contralateral squares fixed, we trained a linear support vector machine to discriminate the color of a single contralateral square based on activity during the sample from units for which we recorded at least 14 trials

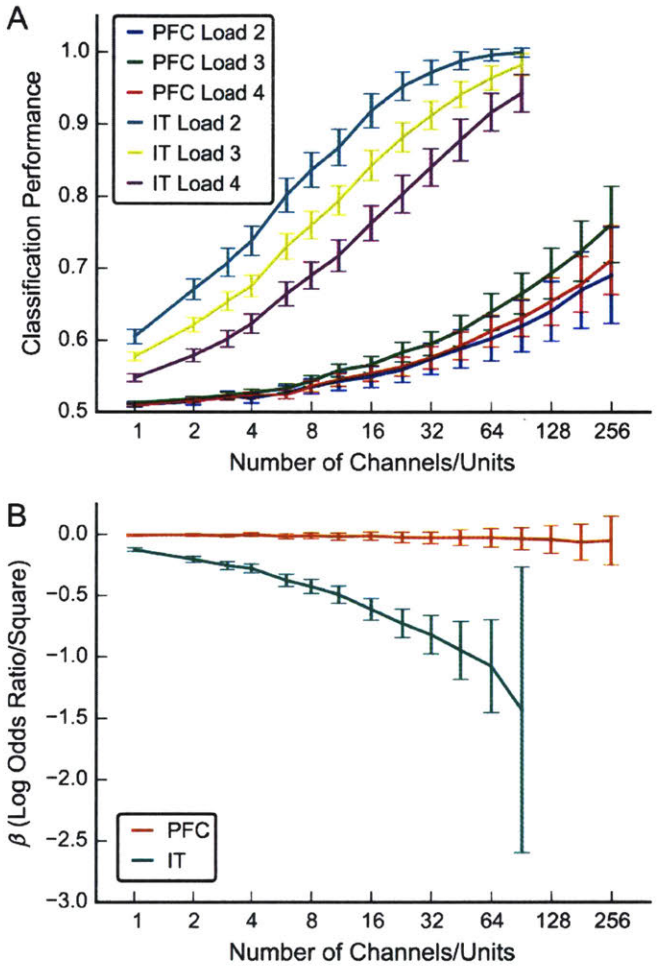


Figure 3-9: Change in Color Classification Performance with Load. **A**: Classification performance by load condition and number of units, in PFC ($n = 361$) and IT ($n = 127$). **B**: Change in log odds ratio of correct classification with the number of squares, in PFC and IT. Error bars are the subsampling estimate of SEM with finite population correction (see Methods).

of each contralateral condition. We trained the classifier on all but one trial and tested its performance on the held out trial, for all 14 possible train/test splits. We then determined how classification performance changes with the number of squares, using logistic regression to determine how the log odds ratio of correctly classifying the color of a given square changed with the number of squares present in the display (see Methods).

This analysis measures information about an individual square, rather than all squares, but the results correspond to expectations from the analysis of explained variance above. In IT, classification accuracy using random subsamples of 1 to 64 units decreased as load increased ($p < 0.005$, test of difference in logistic regression β from zero by subsampling). Thus, IT conveyed less information about any given square as load increased, as expected if total information (about all squares) did not increase or increased sublinearly with load. However, in PFC, classification performance did not change with load regardless of the number of channels used (all $p > 0.4$). Thus, PFC conveyed approximately the same amount of information about any given square regardless of load, as expected if information increased linearly with load. Averaged across loads, classification performance based on a single IT unit was equivalent to between 23 and 32 PFC channels (1 IT unit: 57.6%; 23 PFC channels: 56.2%; 32 PFC channels; 58.1%). With these numbers of units, the decrease in log odds ratio of correct classification with load was significantly higher in IT than PFC (1 IT unit vs. 23 PFC channels, $p = 0.02$; 1 IT unit vs. 32 PFC channels, $p = 0.04$, subsampling), confirming that IT was more strongly capacity-limited than PFC.

3.4.4 Effect of Crowding

As shown in Figures 3-4 and 3-5, PFC was more sensitive to position than IT. Thus, one possible explanation for the difference between effects of load in PFC and IT is a difference in receptive field sizes. If neurons only represent the contents of their receptive fields but have a fixed representational capacity, then presenting additional squares outside of the receptive fields would increase the amount of information represented at the population level. Presenting additional squares inside of the receptive fields would leave the amount of

information unchanged, or perhaps produce a decrease due to interference.

To examine this possibility, we manipulated the distance between squares, so that two squares were either "crowded" or "uncrowded." In both uncrowded and crowded conditions, we presented one square in either the upper or lower quarter-field, with the quarter-field randomized by session. In uncrowded conditions, we presented a second square along the horizontal meridian. In crowded conditions, we presented the second square adjacent to the first (Figure 3-10A). This second square had the same pair of possible colors in both conditions. We computed ε_p^2 for the 4 possible color combinations of these two squares, and compared values between the crowded and uncrowded conditions at loads 3 and 4.

Color selectivity in PFC did not decrease with crowding. Instead, at sample and during the delay, the square colors explained more variance in the crowded condition compared to the uncrowded condition (sample: $F(1,366) = 10.6$, $p = 0.001$; delay: $F(1,366) = 7.0$, $p = 0.009$, ANOVA on condition and load with repeated measure of unit; Figure 3-10B). Explained variance (ε_p^2) was greater in the crowded condition at both load 3 (sample: $t(366) = 2.6$, $p = 0.01$; delay: $t(366) = 2.1$, $p = 0.04$, t-test) and load 4 (sample: $t(366) = 2.4$, $p = 0.01$; delay: $t(366) = 2.3$, $p = 0.02$). In IT, there was no significant effect of crowding ($F(1,154) = 1.5$, $p = 0.23$, ANOVA). Although we are unsure why color selectivity in PFC should increase with crowding, the presence of an increase, rather than a decrease, indicates that differences in modulation by load between PFC and IT cannot be explained solely by differences in receptive field sizes.

We performed further analyses to examine possible explanations for the increase in information in PFC in the crowded conditions. We first examined whether the increase result from systematic bias in color selectivity for the crowded square positions compared to the horizontal meridian. At load 2, there was no significant difference in the proportion of variance explained by square color between the two positions in either region (PFC sample: $t(366) = -0.78$, $p = 0.44$; PFC delay: $t(366) = -0.35$, $p = 0.73$; IT: $t(154) = -0.41$, $p = 0.68$). We next apportioned the variance explained by color into linear and nonlinear components (Figure 3-10C). The linear component reflects variance explained by modeling

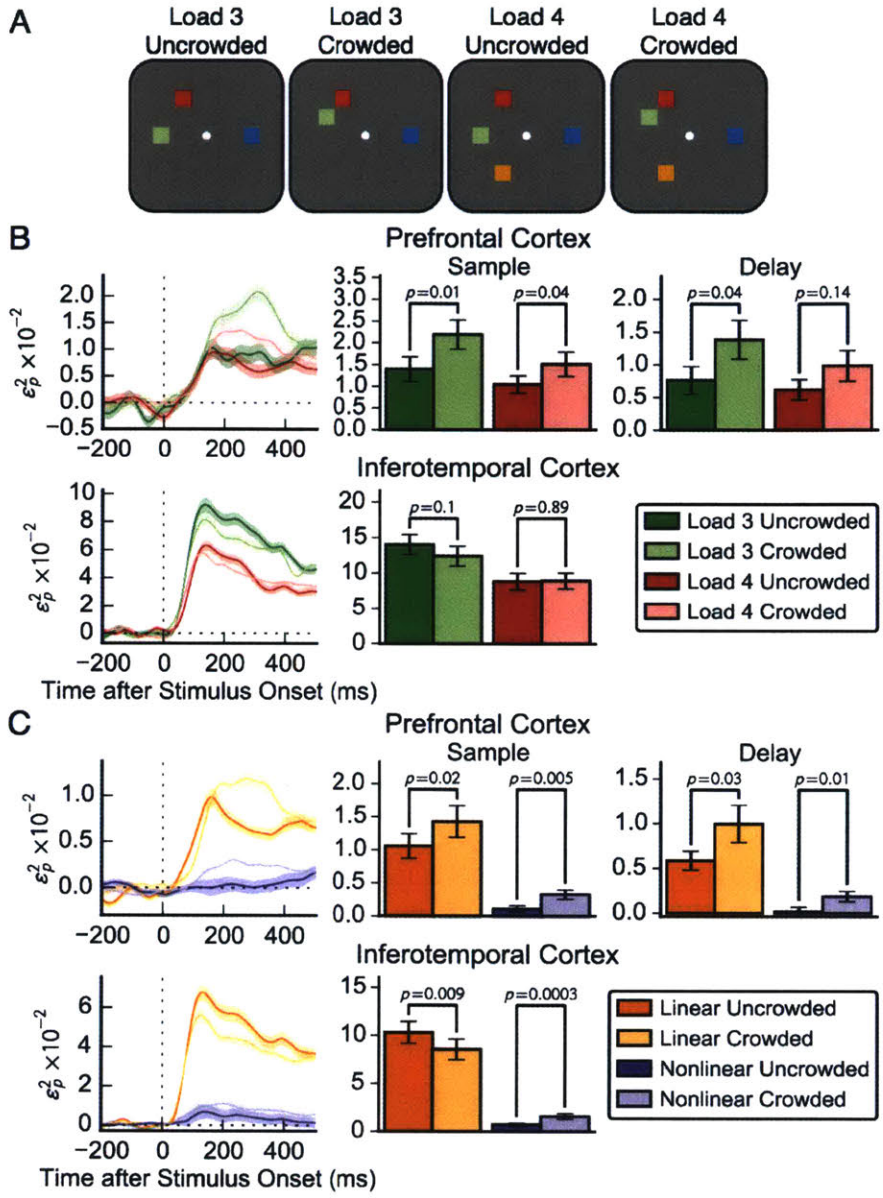


Figure 3-10: Changes in Color Selectivity with Crowding. **A**: Uncrowded and crowded conditions. In the uncrowded condition, one square was presented along the horizontal meridian in the contralateral hemifield. In the crowded condition, this square was instead presented at 45° from the horizontal meridian, separated by 2.2 degrees of visual angle from another square. **B**: Debiased proportion of variance explained (ε_p^2) by the colors of two uncrowded or crowded squares in PFC and IT, in uncrowded (dark colors) and uncrowded (light colors) conditions at loads 3 (green) and 4 (red). **C**: Debiased proportion of variance explained by linear (orange) and nonlinear (purple) interactions between the two squares, in uncrowded (saturated colors) and crowded (unsaturated colors) conditions. p-values reflect significance by a dependent samples t-test. Error bars are ± SEM.

firing rates as the sum of coefficients reflecting the square colors; the nonlinear component reflects variance explained by the specific combination of the two colors. In PFC, we found that both linear and nonlinear explained variance increased in the crowded condition compared to the uncrowded condition during the sample and delay (linear sample: $t(366) = 2.3$, $p = 0.02$; nonlinear sample: $t(366) = 2.6$, $p = 0.005$; linear delay: $t(366) = 2.2$, $p = 0.03$; nonlinear delay: $t(366) = 2.5$, $p = 0.01$), but there was no significant interaction (sample: $F(1, 366) = 0.76$, $p = 0.38$; delay: $F(1, 366) = 1.6$, $p = 0.21$). However, in IT, linear information decreased with crowding ($t(154) = 2.6$, $p = 0.009$), while nonlinear information increased ($t(154) = 3.7$, $p = 0.0003$), and the interaction was significant ($F(1, 154) = 13.1$, $p = 0.0004$). Thus, although crowding did not significantly affect the overall proportion of variance explained by color in IT, it did affect how color was represented.

3.4.5 Effects at Test

When the test screen is presented, the animal must use its knowledge of the previously presented stimuli to make a saccade. We investigated how the changed square is represented during this epoch. We fit a model including the test display, sample display, and the square that changed, and tested how much variance each could explain when excluding contributions of the remaining factors. Figure 3-11A shows the results. In the window from 70 to 161 ms after test onset, activity in both PFC and IT reflected all three factors (all $p < 10^{-4}$, permutation test; see Methods). PFC also showed weak but significant color-selective persistent activity during the final 91 ms of the delay ($\varepsilon_p^2 = 0.0036$, $p < 10^{-4}$), and variance explained by the sample color did not change significantly at test ($t(366) = -0.5$, $p = 0.60$, t-test). By contrast, IT showed no significant color-selective persistent activity at the end of the delay ($\varepsilon_p^2 = -0.0006$, $p = 0.79$, permutation test), but carried sample color information at test (difference: $t(154) = 4.3$, $p < 10^{-6}$, t-test). Sample color explained more variance at test in IT than PFC ($t(227.8) = 3.9$, $p = 0.0001$, unequal variance t-test), and the difference between variance explained by sample color during the delay versus test was significantly greater in IT than PFC ($t(235) = 5.0$, $p < 10^{-5}$).

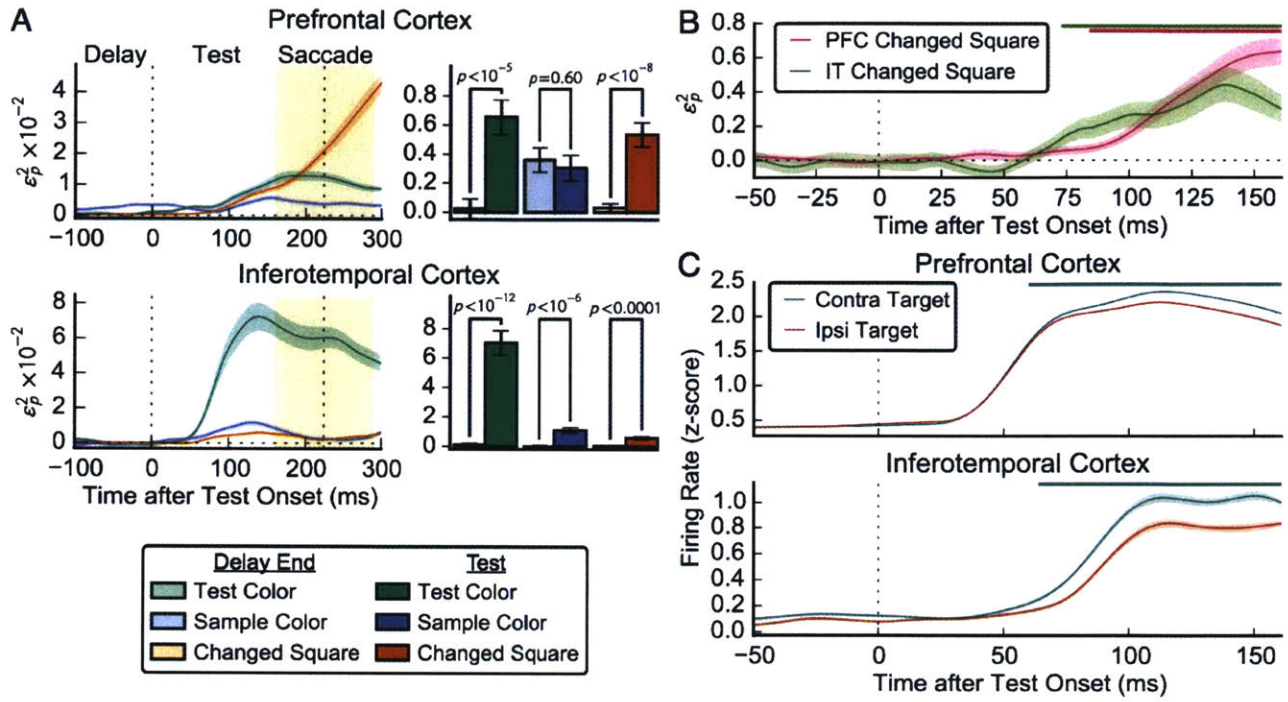


Figure 3-11: Selectivity at Test. **A**: Left: Debiased proportion of variance explained (ε_p^2) by test color, sample color, and changed square over time. Each proportion excludes variance explainable by the remaining factors. Dotted vertical lines indicate the onset of the test stimulus and the median saccade time. Goldenrod shaded region indicates the central 90% of the saccade distribution. Firing rates were smoothed with a Gaussian of 75 ms FWHM. Right: Debiased proportion of variance explained by each factor, compared between the final 91 ms of the delay period (light colors) and the period from 70 ms after test onset until the 5% quantile of the saccade latency at 161 ms (dark colors). Error bars are ± SEM. **B**: Comparison of the time course of ε_p^2 for the changed square in PFC and IT, before the saccade. Bars at top reflect significant time points at $\alpha = 0.05$ (permutation test, corrected for tests at time points after test stimulus onset; see Methods). Error bars are ± SEM. **C**: Comparison of firing rates between trials when the changed square was contralateral or ipsilateral to the recording site, averaged within test displays. Firing rates were baseline-subtracted and normalized by the baseline standard deviation. Bars at top reflect significant time points at $\alpha = 0.05$ (one-sample t-test, Benjamini-Hochberg corrected). Error bars are ± SEM after removing unit means. In **A**, firing rates were smoothed with a Gaussian of 75 ms FWHM; in **B** and **C**, they were smoothed with a raised cosine function of 20 ms FWHM.

Since the task required that the animal make a saccade to the location of the changed square, we investigated the time course with which the two regions signaled this square's location. Figure 3-11B shows the time course of explained variance by the changed square location in IT and PFC. Explained variance exceeded that of a permutation distribution 74 ms after stimulus onset in IT, versus 85 ms after stimulus onset in PFC; the difference was not statistically significant ($p = 0.70$, bootstrap test). We also investigated firing rates directly, taking the mean across each test stimulus display separately for trials when the target was the ipsilateral square and trials when the the target was one of the contralateral squares. In both regions, firing rates were significantly higher for contralateral targets than ipsilateral targets (PFC: $t(366) = 8.3$, $p < 10^{-14}$; IT: $t(153) = 4.2$, $p < 10^{-4}$, t-test on normalized firing rates from 70 to 161 ms after test onset). Latencies were comparable. In PFC, the difference was first significant 61 ms after stimulus onset, while in IT, the difference was significant 65 ms after stimulus onset.

We further examined whether either region was related to the animal's behavior. We computed the proportion of variance explained by the changed square position, sample color, and test color separately for correct and incorrect trials at load 3 or 4 (Figure 3-12). We tested for significance by permuting between groups while holding the load condition constant, and comparing the true value at each time point against the maxima or minima of the permutations across all time points (see Methods). In PFC, a significantly greater proportion of variance was explained by the changed square position on correct trials compared to incorrect trials ($p < 0.002$, two-sided permutation test). This difference was only significant starting at 131 ms after test stimulus onset, and may thus be related to saccade preparation, rather than stimulus processing. In IT, a significantly greater proportion of variance was explained by the sample color on correct trials, from 91 and 119 ms after test stimulus onset ($p = 0.006$). Variance explained by the changed square position did not differ on incorrect trials in IT ($p = 0.16$), and variance explained by the sample color did not differ significantly in PFC ($p = 0.16$). Moreover, in both regions, variance explained by test color did not differ significantly on incorrect trials (PFC: $p = 0.79$, IT: $p = 0.07$), and there was no significant

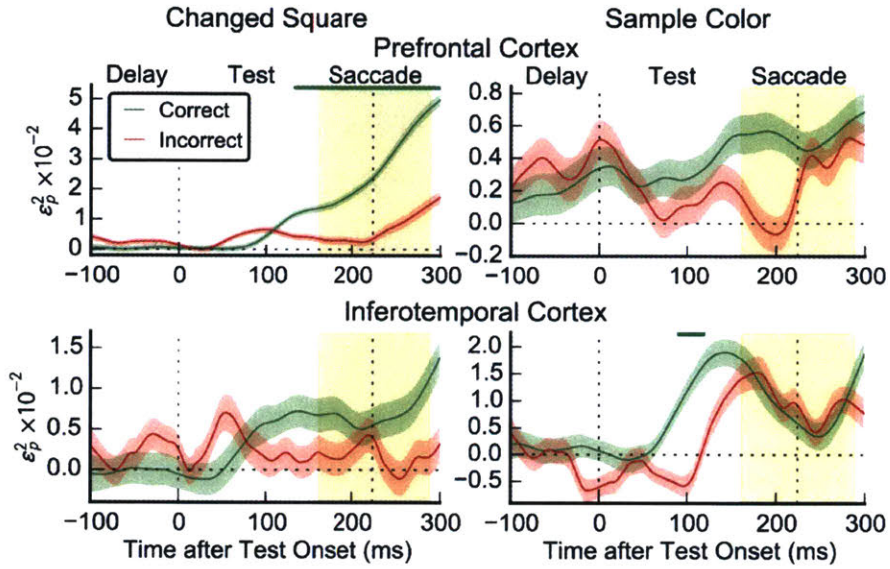


Figure 3-12: Differences Between Correct and Incorrect Trials at Test. Comparison of variance explained by the changed square location (left) and sample color (right) in PFC (top) and IT (bottom) on correct trials (green) and incorrect trials (red). Dotted vertical lines indicate the onset of the test stimulus and the median saccade time. Goldenrod shaded region indicates the central 90% of the saccade distribution. Firing rates were smoothed with a Gaussian of 75 ms FWHM. Green bars at tops of plots indicate time points where significantly greater variance was explained on correct versus incorrect trials at $\alpha = 0.05$ (two-sided permutation test, corrected for tests at every time point; see Methods). Error bars are $\pm$ SEM after removing average unit-level effects.

difference in variance explained by sample color at sample presentation or during the delay (PFC sample: $p = 0.46$, IT sample: $p = 0.37$, PFC delay: $p = 0.13$, IT delay: $p = 0.14$).

3.5 Discussion

Our study reveals three dissociations between PFC and IT during visual short-term memory for multiple items. First, during the sample period of the change localization task, activity in PFC primarily reflected square positions, while activity in IT primarily reflected square colors, although both areas encoded both attributes. Second, the representational capacity of PFC for square colors increased with load, while the representation capacity of IT remained constant or decreased. Finally, when the test screen was presented, both PFC and IT represented which square had changed, with similar time courses. However, the nature of this representation differed between regions. In IT, selectivity for the changed square location arose simultaneously with selectivity for the sample colors, which was not present at the end of the delay. The strength of sample color selectivity, but not changed square location selectivity, correlated with behavioral performance. By contrast, in PFC, selectivity for sample colors was consistently present throughout the delay and test periods. The strength of changed square location selectivity, but not sample color selectivity, correlated with behavioral performance.

While previous studies did not explicitly compare position and color selectivity between PFC and IT, our results are broadly consistent with earlier reports. Both PFC and IT neurons are known to be spatially selective (Rao, Rainer, & Miller, 1997; DiCarlo & Maunsell, 2003; Meyer, Qi, Stanford, & Constantinidis, 2011; Kadohisa et al., 2015; Hong, Yamins, Majaj, & DiCarlo, 2016) and possess spatial receptive fields (Suzuki & Azuma, 1983; Rainer, Asaad, & Miller, 1998; Op De Beeck & Vogels, 2000; Bichot et al., 2015; Riley, Qi, & Constantinidis, 2016). In both regions, receptive fields are predominantly contralateral (Rainer et al., 1998; Op De Beeck & Vogels, 2000), consistent with our finding that neither region contained a strong representation of ipsilateral stimuli. Two studies have explicitly inves-

tigated color selectivity in PFC during multiple-item visual short-term memory. Like our study, Lara and Wallis (2014) reported substantially stronger position selectivity than color selectivity in PFC. Although they found no significant delay period color selectivity whatsoever, in our data, delay period color selectivity was present but weak. This discrepancy may result from differences in recording sites: We found the strongest concentration of sites yielding color selectivity in the ventral prearcuate region, which they did not sample. By contrast, Buschman et al. (2011) found marginally stronger color selectivity in PFC than we report. However, they restricted their analysis to selective units. In IT, we targeted a region known for its color selectivity (Komatsu, Ideura, Kaji, & Yamane, 1992; Harada et al., 2009; Lafer-Sousa & Conway, 2013), and the color selectivity we observe is in line with these previous studies.

Our observations regarding changes in color information with load in PFC differ slightly from previous literature. Buschman et al. (2011) reported impaired encoding of individual squares as the number of squares presented increased. In this study, we found that total variance explained by the colors of all squares doubled as the number of contralateral squares increased from two to three, although there was no significant increase in explained variance from three squares to four. Additionally, we found no dependence of decoding accuracy on contralateral load. This discrepancy likely results from differences between our analysis and theirs. When measuring proportion of variance explained, we included not only linear effects of color, but also their interactions. Additionally, our decoding analysis measured classification performance while holding colors of other squares constant. Thus, our measures of color selectivity incorporate both linear effects of individual square colors and nonlinear interactions between squares. Since Buschman et al. (2011) computed variance explained by square color without taking into account colors of other squares, they measured only linear effects of individual square colors, and not nonlinear interactions between squares. Such nonlinear interactions are a form of mixed selectivity, an inherent and behaviorally relevant feature of prefrontal neurons (Rigotti et al., 2013; Fusi, Miller, & Rigotti, 2016). Applying our analysis to their data yielded results qualitatively similar to those in Figure 3-7. The

absence of a detectable capacity limit in PFC is in line with the lack of saturation with capacity in measures based on local field potentials (Kornblith, Buschman, & Miller, 2016).

Previous studies have reported attenuated activity when multiple stimuli are presented simultaneously (Sato, 1989; E. K. Miller, Gochin, & Gross, 1993; Rolls & Tovee, 1995; Chelazzi et al., 1998; Missal et al., 1999; Zoccolan et al., 2005). Since responses to simultaneously presented stimuli closely resemble the average response to the individual stimuli, this attenuation is believed to result from a normalization process (Zoccolan et al., 2005). Consistent with these previous studies, we found that manipulations of stimulus load or position had little effect upon the total proportion of variance in IT firing rates that could be explained by the stimulus colors.

Beyond stimulus representation at sample presentation, our study provides insight into how the animal might maintain these stimuli in short-term memory and determine which stimulus changed at test. Although LIP and PFC exhibited robust stimulus-selective delay period activity, as at sample presentation, they primarily represented the positions of the stimuli, and not their colors. In IT, we observed very little stimulus-selective delay period activity, and the population ceased to reflect sample stimulus colors prior to test stimulus onset. Firing rates were suppressed relative to baseline independent of the stimuli to be maintained. Nonetheless, we observed a resurgence of modulation by sample color as neurons responded to the test stimulus. In PFC, modulation by sample color did not change between the end of the delay and the test period, but shortly after test stimulus onset, both PFC and IT signaled which stimulus in the display had changed.

Compared with previous studies of short-term memory in IT, we found weaker stimulus-selective delay period activity, but comparable modulation by stimuli held in short-term memory at test. In delayed match to sample and visual search tasks, stimulus-selective delay period activity is present in 10-50% of IT units (Fuster & Jervey, 1982; Miyashita & Chang, 1988; Chelazzi et al., 1998; Freedman, Riesenhuber, Poggio, & Miller, 2003; Sobotka, Diltz, & Ringo, 2005; Woloszyn & Sheinberg, 2009). We observed stimulus-selective delay period activity in only 3% of recorded units (4/155). As we discuss in Appendix A, this

discrepancy is unlikely to arise from differences in statistical power. It may instead relate to differences between tasks: While most previous studies rewarded animals for identifying matching stimuli, we rewarded animals for identifying the stimulus that did not match. But even in the near absence of stimulus selectivity during the delay, we observed modulation by the sample stimulus at test stimulus presentation, consistent with previous reports (E. K. Miller, Li, & Desimone, 1993; E. K. Miller & Desimone, 1994; Chelazzi et al., 1998; Pagan, Urban, Wohl, & Rust, 2013). This modulation reflected both the location of the changed square and the direction of the color change, distinguishing a change from color 1 to color 2 from a change from color 2 to color 1.

How could IT neurons signal a color change without persistent activity? A common hypothesis in the literature is that such signals originate from PFC (E. K. Miller et al., 1996; Chelazzi et al., 1998; Pagan et al., 2013). However, in our task, color representation was weak in PFC, both at sample and during the delay. We cannot dissociate selectivity to a specific color change from a combination of persistent activity reflecting the sample stimulus and a visual response to the test stimulus. However, we observed no increase in selectivity for the sample stimulus at test stimulus presentation relative to the preceding delay. Such an increase might be expected if PFC provided a feedback signal to IT reflecting the sample colors only after the test stimulus was presented. An alternative possibility is that IT represents the previously presented stimulus not through persistent activity, but through changes in synaptic weights. Whether such changes in synaptic weights truly constitute "working memory" is unclear. In PFC, such synaptic changes have been termed "activity-silent working memory" (Stokes, 2015), and modeling studies have confirmed their adequacy to support behavior in working memory tasks (Mongillo et al., 2008; Lundqvist et al., 2011). However, IT neurons show similar dependence of visual response upon stimulus history in the absence of a task (Vogels, Sáry, & Orban, 1995) and even under anesthesia (E. K. Miller, Gochin, & Gross, 1991). These adaptation effects typically take the form of repetition suppression (E. K. Miller et al., 1991; Vogels et al., 1995; Sawamura, Orban, & Vogels, 2006; Vogels, 2016).

Given that IT can signal stimulus changes even in the absence of a task, are such signals behaviorally relevant? An animal could conceivably implement our change localization task by simply averaging activity in IT related to the square locations, and saccading to the location producing the strongest response. We found a link between the strength of the representation of the sample stimulus at test stimulus presentation and the animal's performance, suggesting that a causal role is plausible. Moreover, previous studies have implicated such passive signals in visual short-term memory tasks. E. K. Miller and Desimone (1994) showed that animals trained to perform a delayed match to sample task responded to any repeated stimulus as a match, even if it did not match the initial sample. Only after subsequent training with repeated stimuli did animals learn to compare stimuli to the sample. Most neurons were suppressed by match stimuli; these neurons were also suppressed by repeated non-match stimuli. Some neurons showed enhanced responses to match stimuli; these neurons did not show enhanced responses to repeated non-match stimuli. However, the proportion of enhanced neurons increased after training with repeated non-match stimuli, and in one of the two animals, no neurons showed match enhancement prior to training. Thus, the default strategy that monkeys use to perform short-term memory tasks may rely on passive repetition suppression effects in IT. Stimulation studies provide additional evidence for a role of IT in detecting stimulus changes. When Parvizi et al. (2012) stimulated a face-selective region in human temporal cortex while the subject attended to non-face objects, the subject reported a non-specific sensation that something in his visual field had changed. In monkeys, stimulation of anterior IT sites at sample presentation results in little deficit in delayed match to sample tasks, and evidence for deficits from delay period stimulation is mixed, but stimulation during test stimulus presentation produces profound deficits (Kovner & Stamm, 1972; Delacour, 1977; Moeller, Crapse, Chang, & Tsao, 2017).

What, then, is the role of PFC and LIP in our task? Although both regions only weakly represented square colors, they robustly represented stimulus positions, and this robust representation persisted throughout the delay. To some degree, this position-selective persistent activity may be automatic: It is present even before training on a visual short-term mem-

ory task (Meyer et al., 2011; Meyers, Qi, & Constantinidis, 2012). However, it could also play a more active role, helping to direct attention at test stimulus presentation and shield stimulus representations in visual cortex from interference. Human studies suggest that directing attention toward a particular stimulus location long after sample stimulus offset, but before test stimulus onset, greatly enhances behavioral performance (Griffin & Nobre, 2003; Landman, Spekreijse, & Lamme, 2003; Makovski, Sussman, & Jiang, 2008). These effects are present even when the test display probes only a single item (Griffin & Nobre, 2003), or when compared to a condition in which a location cue appears simultaneously with the test display (Landman et al., 2003; Makovski et al., 2008). Given that adaptation effects in IT dissipate shortly after the onset of the response to the stimulus (Vogels, 2016), maintaining a representation of stimulus positions may prepare LIP and PFC to process the test stimulus display and convert IT input to a behavioral response. In tasks involving maintenance of color information, the relationship between activity in human parietal cortex and individual differences in visual short-term memory capacity (Vogel & Machizawa, 2004; Vogel et al., 2005; Todd & Marois, 2005) may reflect the importance of stimulus position signals in the readout of stimulus changes, rather than a direct role for parietal cortex in representing the maintained stimuli.

In summary, our study directly implicates inferotemporal cortex as a bottleneck in visual capacity, and in performance of a multiple-item visual short-term memory task. Compared with prefrontal cortex, selectivity for colors of visual stimuli in IT was both stronger and more strongly affected by load. Although persistent activity in both PFC and IT barely reflected colors of previously presented stimuli across the delay, IT activity reflected the difference between sample and test stimuli, and the strength of this representation predicted performance. While activity in IT reflects stimulus changes even in the absence of persistent activity and cognitive demand, it may nonetheless play an essential role in visual short-term memory tasks.

3.6 Contributions and Acknowledgements

I designed the experiments in this chapter with Earl K. Miller, and performed the experiments with Roman Loonis, Jefferson E. Roy, and Jacob A. Donoghue. I analyzed the data and wrote the chapter. All authors commented on the manuscript.

We thank S. L. Brincat for comments on the manuscript. This work was supported by an NIMH Conte Center grant, The Picower Foundation, and an NSF Graduate Research Fellowship.

Chapter 4

Conclusion

4.1 Summary of Results

Our studies investigated how the number of items to be maintained in visual short-term memory affect local field potentials and representations throughout the brain. Table 4.1 summarizes the main findings. Both experiments examined neural activity recorded over the course of a change localization task, in which animals viewed a sample display consisting of a set of colored squares, remembered them across a brief delay, viewed a subsequent test display with a change to the color of one of the squares, and made a saccade to this changed square to receive juice reward. As the number of items to be remembered increased, animals' performance was substantially impaired.

In chapter 2, we showed that this behavioral limit to cognitive capacity is not manifested in local field potentials recorded from the primate lateral intraparietal area (LIP), frontal eye field (FEF), and lateral prefrontal cortex (PFC) in any recognizable way. In both regions, we observed changes in oscillatory power and synchrony that correlated with the number of items to be remembered, but did not saturate, even when performance was heavily degraded. Nonetheless, there were clear dissociations between effects at sample presentation and during the delay, and between power changes at lower (alpha, beta, and low gamma) frequencies and changes at higher (high gamma) frequencies. As monkeys viewed the stimulus array,

Finding	Figure
Low frequency power at stimulus presentation tracks total load, whereas high frequency power primarily tracks contralateral load.	2-4
Low frequency power decreases with load at stimulus presentation, but increases with load during the delay.	2-4
Oscillatory power is modulated by load beyond behavioral capacity.	2-5
Although low frequency power is modulated by the total number of stimuli present, position information in power primarily reflects contralateral stimulus positions.	2-6
Low frequency power decreases are similarly strong in position-selective and non-position-selective electrodes.	2-7
Position and color information in PFC are strongest within the same circumscribed region.	3-3
LIP and PFC primarily represent stimulus positions, whereas IT primarily represents stimulus colors.	3-4, 3-5
Stimulus load has little effect upon overall firing rates in LIP, PFC, or IT.	3-6
IT has a stronger limit to representational capacity for stimulus colors than PFC.	3-7, 3-8, 3-9
IT represents stimulus changes in the near-total absence of stimulus-selective persistent activity.	3-11
The strength of IT representations of stimulus changes correlates with behavioral performance.	3-12

Table 4.1: Main Findings.

lower frequency power decreased with the total number of stimuli present, while higher frequency power instead increased, and this increase was driven primarily by the number of contralateral stimuli. As the animals maintained the previously presented stimuli in visual short-term memory, effects at lower frequencies was reversed: Power increased with load in proportion to the number of contralateral items. No effects were present at higher frequencies. We further showed that the power decrease at lower frequencies during sample presentation was present in all electrodes regardless of position selectivity, while the power increases at higher frequencies during sample presentation and at lower frequencies during the delay was primarily localized to electrodes that conveyed information regarding stimulus positions.

In chapter 3, we investigated stimulus representation in a similar task in LIP, PFC, and inferotemporal cortex (IT), which provides the predominant ventral stream input to PFC. We found striking dissociations between these regions. Although square positions were relevant during the test period of the task, the task required only that colors, and not positions, were maintained across the delay period. Nonetheless, at both sample and delay, LIP and PFC

preferentially represented square positions, rather than colors. By contrast, at sample, IT preferentially represented square colors, rather than positions. However, all three regions conveyed some information regarding both stimulus position and color, and we were able to measure changes in color representation in PFC and IT as load increased. We found that PFC conveyed more total color information when more colored squares were present on the screen, in line with our finding that local field potential measures did not saturate, while IT conveyed the same amount of color information or less. Thus, the representational capacity of IT, but not PFC, appears to be fixed. While PFC and LIP conveyed substantial information regarding square positions as the animal maintained the previously presented stimulus in memory, little color selectivity was present in any recorded region across this brief delay. However, in IT, but not PFC, we observed a resurgence of information at test presentation, reflecting the difference between the sample and test displays. Both regions signaled the location of the changed square with similar, short latencies. Activity at test presentation in both regions was related to behavior, but in different ways. PFC more strongly represented the location of the changed square on correct trials than incorrect trials, but this difference had a relatively long latency, and may reflect its role in generating the saccade to the changed square. By contrast, IT more strongly represented the difference between sample and test square colors on correct trials, with a latency consistent with the onset latency of the visual response. This study implies a stronger capacity limit in IT than PFC, and suggests that IT may directly participate in producing the behavioral response.

4.2 Non-Selective Modulation in Cognitive Tasks

These findings demonstrate striking dissociations between the degree of neural activity and the information represented in these measures. In chapter 2, we found that the power of lower frequency oscillations at stimulus presentation scaled with the number of items displayed on the screen, and these changes were equally strong regardless of the hemifield in which the items were presented (Figure 2-4). The same frequency band also represented

stimulus positions, but this representation was far stronger for items presented in the hemifield contralateral to the recording site than for ipsilaterally presented items (Figure 2-6), and modulation by load was present even in electrodes that did not carry position information (Figure 2-7). Reexamination of the data from (Buschman et al., 2011) revealed a substantially stronger representation of contralateral square positions and colors than ipsilateral square positions and colors in PFC spiking (position: $\varepsilon^2 = 0.040$ vs. 0.0081, $t(574) = 9.0$, $p < 10^{-17}$, t-test; color: $\varepsilon^2 = 0.0066$ vs. 0.0011, $t(574) = 4.8$, $p < 10^{-5}$). Thus, these lower frequency oscillations changed with ipsilateral load, but changes did not track information represented within the region, either in spiking or in the local field potentials themselves.

In chapter 3, LIP, PFC, and IT all showed similarly elevated firing rates at the end of the delay of the change localization task, as the animal anticipated the onset of the test stimulus (Figure 3-6). This "ramping" in firing rate toward the end of the delay is a common feature of neural activity during cognitive tasks (Rainer & Miller, 2002; Brody, Hernández, Zainos, & Romo, 2003). In prefrontal cortex, it is often accompanied by increased selectivity for the stimulus held in visual short-term memory. However, in IT, we observed no selectivity for previously presented stimuli during this period. Importantly, this ramping does not appear to result purely from decay of recurrent activity following visual stimulation. Although we do not present the results in chapter 3, we also recorded from the same regions in the same sessions during a delayed saccade task with a shorter delay (1 second vs. 1.2 to 1.5 seconds in the change localization task). In this task, ramping was also present in all three regions, and occurred earlier relative to the change localization task. Ramping thus appears to reflect anticipation of the end of the delay period, rather than the time elapsed since stimulus presentation.

These findings reveal the extent of non-selective modulation throughout the brain, and underscore the difficulty of understanding the brain's representations based on net measures of neural activity. We remain unsure of the sources of these non-selective signals. The low frequency oscillations may reflect integration of intrinsic input and input from contralateral PFC among circuits generating beta oscillations. The ramping signals in IT may originate

from LIP or PFC, while the ramping signals in the latter regions may be generated intrinsically. Alternatively, either form of non-selective modulation could originate or propagate through afferents from subcortical regions such as the thalamus or midbrain.

4.3 What Does Prefrontal Cortex Represent?

Much of the literature regarding primate lateral prefrontal cortex emphasizes its flexibility (E. K. Miller & Cohen, 2001; Barbas & Zikopoulos, 2007; Stokes et al., 2013). However, the results in chapter 3 suggest that prefrontal representations are surprisingly inflexible. In a task requiring active maintenance of colors in visual short-term memory, most selective prefrontal neurons represented only stimulus positions. Moreover, comparison of Figures 3-4 and 3-5 reveals strikingly little difference in prefrontal representation between passive fixation and the task. In Appendix B, we show further results obtained from PFC during passive fixation, revealing the robustness of PFC activity in the absence of a task.

While most studies of PFC have investigated activity as animals performed a cognitive task, studies performed during passive viewing show many of the same characteristics. An early study showed that lateral PFC neurons have circumscribed receptive fields that can be mapped in a passively fixating animal (Suzuki & Azuma, 1983). Scalaidhe, Wilson, and Goldman-Rakic (1999) localized a subregion of PFC containing a high proportion of neurons that responded selectively to faces during passive fixation. Many such neurons exhibited prolonged responses for up to 200 to 1500 ms after stimulus offset, even in an untrained animals. Meyers et al. (2012) recorded from PFC before and after animals were trained on a delayed matching task. Activity during passive viewing prior to training closely resembled activity during task execution after training. McKee, Riesenhuber, Miller, and Freedman (2014) found marginally greater category selectivity as animals performed a task requiring categorization of morphed cat and dog stimuli than when animals passively viewed the stimuli, although a classifier was capable of decoding the categories from population activity in both cases.

Existing literature provides little insight into the nature of PFC stimulus representation. Categorization tasks indicate that, at least for highly trained stimuli, PFC neurons can be sensitive to high-level, behaviorally relevant features (Freedman, Riesenhuber, Poggio, & Miller, 2001; Freedman et al., 2003; McKee et al., 2014). Several studies have compared the selectivity of LPFC neurons for stimulus location and identity (Everling, Tinsley, Gaffan, & Duncan, 2006; Meyer et al., 2011; Kadohisa et al., 2015; Bichot et al., 2015; Riley et al., 2016). These studies suggest that selectivitiy for stimulus location is generally stronger and more prevalent, as we found in chapter 3. The extent of the difference in selectivity for stimulus positions versus features in PFC versus IT is somewhat surprising, given that ventrolateral PFC anterior to the arcuate is believed to receive most of its visual input from IT (Saleem et al., 2014). It is possible that the representation of position is inherited from FEF, which is more strongly connected with retinotopic visual areas (Stanton, Bruce, & Goldberg, 1995; Schall, Morel, King, & Bullier, 1995) and is known to contain a large population of neurons with localized visuospatial receptive fields (Mohler, Goldberg, & Wurtz, 1973). Alternatively, the spatial selectivity we observe in PFC may arise from inputs from the posterior portion of IT, where neurons exhibit greater retinotopic organization and selectivity to spatial position than in anterior portion from which we recorded in chapter 3 (Boussaoud, Desimone, & Ungerleider, 1991; Yasuda, Banno, & Komatsu, 2010).

The studies presented in chapters 2 and 3 use artificial stimuli (colored squares), and may not generalize to naturalistic stimuli or scenes. Given its position at the apex of the cognitive hierarchy, PFC visual representations may be selectively tuned to complex, ecologically relevant stimuli. Alternatively, PFC may mix representations of differing levels of sophistication originating from many levels of the visual system. Since most cognitive tasks involve the discrimination or maintenance of visual stimuli, understanding what kinds of visual information enter PFC could yield novel insight into how these representations are transformed to produce behavior. Given the ease with which PFC representations can be activated during passive visual paradigms, the same approaches to understanding visual representation that have proved fruitful in IT can be easily applied to PFC.

4.4 Parallel Mechanisms for Short-Term Memory Revisited

On the basis of differences between the characteristics of populations of IT neurons that were suppressed or enhanced to matching stimuli relative to non-matching stimuli, E. K. Miller and Desimone (1994) proposed that short-term memory reflects two parallel mechanisms: One that is passive and automatic, and one that requires active maintenance. Our findings in chapter 3 implicate the passive mechanism in performance of our change localization task. This evidence is somewhat circumstantial: While we show a relationship between neural activity and behavior that is present in the absence of active maintenance signals, we do not know for certain whether the adaptation effects we observe are necessary or sufficient for task performance. A causal role is difficult to prove experimentally, particularly in our change localization task. Although the biological basis of adaptation phenomena is poorly understood (Vogels, 2016), a hypothetical pharmacological manipulation that affects adaptation while leaving stimulus selectivity unchanged could yield some evidence. In a task requiring the animal to report whether or not the display changed, rather than saccade to a changed square, optogenetic manipulation could also provide some insight. If the animal selects the response based on the extent of adaptation-mediated repetition suppression, then inhibiting IT neurons should bias the animal toward a match response, while activating IT neurons at test stimulus presentation should bias the animal toward a non-match response.

Recent human psychophysics studies have also shown evidence for two separate mechanisms of visual short-term memory. Studies of multiple-item change detection paradigms have compared performance during retro-cued trials, in which a cue signaling a specific item in the display to be examined is presented long after sample offset but prior to test, to post-cued trials, in which the cue is presented simultaneously with the test display. The retro-cue yields a substantial increase in performance relative to the post-cue, effectively doubling visual short-term memory capacity, even if the delay before the cue onset in the retro-cue condition is as long as the delay before the test onset in the post-cue condition (Landman

et al., 2003). Compared to post-cued trials, performance in the presence of a retro-cue is more strongly affected by visual interference, but less strongly affected by cognitive interference. The retro-cue advantage disappears if the stimulus is masked before the retro-cue is presented (Sligte, Scholte, & Lamme, 2008), but retro-cued performance does not change with a concurrent task at encoding, while post-cued performance suffers (Vandenbroucke, Sligte, & Lamme, 2011; but see Makovski, 2012). Transcranial magnetic stimulation of dorsolateral prefrontal cortex impairs performance on post-cued trials, but not retro-cued trials (Sligte, Wokke, Tesselaar, Steven Scholte, & Lamme, 2011), and activity of visual cortex at encoding predicts later performance on retro-cued trials (Sligte, Scholte, & Lamme, 2009). The difference in performance and characteristics between retro-cued and post-cued trials led Sligte et al. (2008) to propose that they reflect two separate visual short-term memory stores: a "fragile" store with higher capacity that can be accessed only with the aid of the retro-cue, and a "robust" store used when the retro-cue is absent.

It is possible that "fragile" visual short-term memory reflects passive storage in visual regions, which is read out by higher regions in the degree of adaptation at test stimulus presentation. In this scenario, the retro-cue improves performance by allowing the subject to direct their attention to the stimulus to be examined prior to test screen onset. Since the effects of adaptation upon visual responses are strongest during the initial response transient (Vogels, 2016), by the time a simultaneously presented cue could be processed, information has already been lost. Because the process is passive, it would be unaffected demands of intervening tasks, but susceptible to distractors, in line with reported findings. While challenging to prove definitively, this hypothesis makes some testable predictions. Behaviorally, a mask presented between the retro-cue and test display should produce a similar performance degradation to a mask presented between the sample display and retro-cue. Electrophysiologically, if the retro-cue exerts its effect through deployment of spatial attention at test stimulus presentation, then attention should modulate the degree of adaptation. To our knowledge, no studies have directly examined adaptation in the context of spatial attention.

Finally, we suggest that task demands and strategies affect which short-term memory

mechanisms are used, particularly for highly trained animals. E. K. Miller and Desimone (1994) showed that training animals to perform a delayed match to sample task with repeated non-match items altered activity in IT, substantially increasing the proportion of neurons that exhibited enhanced responses to a matching stimulus. Vogels and Orban (1994) reported heterogeneity of repetition effects in IT in animals performing delayed match to sample task: In two monkeys, most neurons with significant effects were suppressed to matching stimuli, while in a third, most neurons were enhanced. Although some visual short-term memory tasks, such as those requiring maintenance of serial order (Warden & Miller, 2007, 2010), cannot be solved using adaptation signals alone, many could be, including most delayed match to sample tasks. As we noted in the introduction, human studies find evidence for stimulus-selective persistent activity in visual areas as far back as V1 (Harrison & Tong, 2009; Serences et al., 2009), but monkey studies have reported mixed evidence for such representations in retinotopic cortex (Supèr et al., 2001; Chelazzi et al., 2001; Lee et al., 2005; Hayden & Gallant, 2013), and in chapter 3, we found little evidence for behaviorally relevant stimulus-selective persistent activity in anterior IT. It may be that the brain has multiple mechanisms capable of maintaining the information necessary to solve common short-term memory tasks, and the presence or absence of persistent activity in sensory regions depends on the strategy the subject chooses.

Appendix A

Power Analysis for IT Delay Period Selectivity

In chapter 3, we reported that only 3% of IT units (4/155) were significantly selective for the stimulus display during the delay period of our multiple item visual short-term memory task (permutation test of ϵ^2, $\alpha = 0.01$). Analyzing the power of the study with respect to previous studies is challenging. In addition to the number of stimuli used, number of trials recorded, and statistical test, our ability to detect an effect in a given unit depends on how spiking activity varies across stimuli, trials, and units. Although we use a permutation ANOVA to test for statistical significance, for test against an F distribution with data conforming to the standard ANOVA model (observations normally distributed conditioned on the task condition, with equal within-condition variance), the power for a given unit is ordinarily computed as a function of the asymptotic value of F, the ratio of the true between-condition variance to the true within-condition variance. We may also write F as a function of the

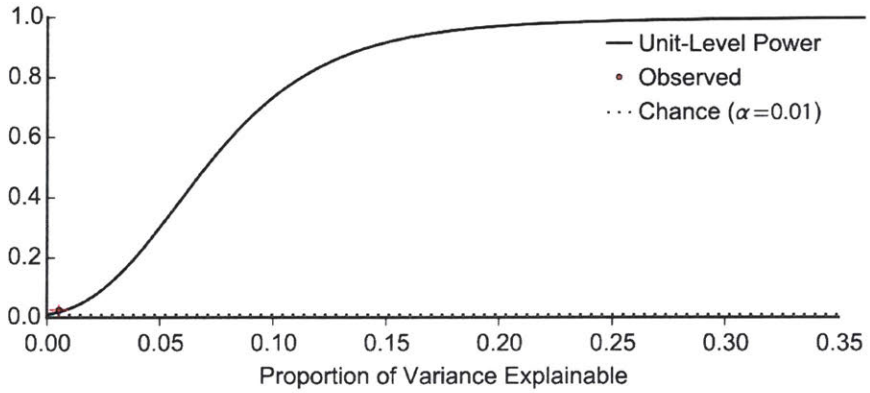


Figure A-1: Power to Detect Significant Units. The solid line represents unit-level power to detect a significant effect of stimulus condition, for different proportions of explainable single-trial variance ($\varepsilon^2_{\text{asymptotic}}$). For tractability, firing rates are assumed to be conditionally normally distributed and homoscedastic, and power is computed based on an F-test rather than the permutation test used in Chapter 3. The orange circle in the lower left reflects the observed mean value of ε^2 (x error bar is 95% CI based on normal approximation) and the observed proportion of selective units (y error bar is 95% Clopper-Pearson binomial CI). The dashed line indicates the α of the test.

asymptotic value of ε^2, the proportion of trial-to-trial variance explainable by the parameters:

$$\begin{aligned}
\varepsilon^2_{\text{asymptotic}} &= \frac{\sigma^2_{\text{between}}}{\sigma^2_{\text{between}} + \sigma^2_{\text{within}}} \\
F_{\text{asymptotic}} &= \frac{\sigma^2_{\text{between}}}{\sigma^2_{\text{within}}} \\
&= \frac{\varepsilon^2_{\text{asymptotic}}}{\varepsilon^2_{\text{asymptotic}} + 1}
\end{aligned}$$

The average unit-level power for different values of $\varepsilon^2_{\text{asymptotic}}$ given the number of trials recorded is shown in Figure A-1. If the value of $\varepsilon^2_{\text{asymptotic}}$ is the same for all units, then the expected proportion of selective units equals the average unit-level power. In general, however, the expected proportion of selective units depends not only upon the mean value of $\varepsilon^2_{\text{asymptotic}}$ but also on its distribution across units. Specifically:

$$P(\text{unit selective}) = \int_0^1 P(\text{unit selective}|\varepsilon^2_{\text{asymptotic}} = x) P(\varepsilon^2_{\text{asymptotic}} = x)\, dx$$

where $P(\text{unit selective}|\varepsilon^2_{\text{asymptotic}} = x)$ is the power function plotted in A-1 and $P(\varepsilon^2_{\text{asymptotic}} = x)$ is the probability density function of $\varepsilon^2_{\text{asymptotic}}$ across units. Unfortunately, no previous studies have attempted to characterize this distribution, and it is likely to depend on the stimuli used in the task.

We also directly compared our results with those of previous studies. Table A.1 shows task parameters and proportions of selective units for previous reports of IT delay period selectivity. Because Woloszyn and Sheinberg (2009) recorded a similar number of trials as we did (when pooling ipsilateral and contralateral stimulus conditions), we chose to compare our results directly using their analysis strategy. We first measured firing rates over the sample period (60-505 ms after stimulus onset), and split them into two halves. We used one half to select the most and least effective displays, and the other to estimate the area under the curve (AUC) and stimulus selectivity index (SSI). As in Woloszyn and Sheinberg (2009), we analyzed only units with AUC > 0.75 and SSI > 0.25 at presentation. For these 31 units, we tested for a significant difference in firing rate over the delay (300-1200 ms after stimulus offset) between trials on which the most effective display was maintained and trials on which the least effective display was maintained. On average, this analysis included 18 trials for each display, the same number as Woloszyn and Sheinberg (2009). However, only 10% of units (3/31) showed differential modulation during the delay (Mann-Whitney U test, $\alpha = 0.05$; $p = 0.20$, binomial test), with a mean AUC of 0.5. This proportion is significantly smaller than the 42% of units (33/79) reported in Woloszyn and Sheinberg (2009) ($p = 0.002$, Fisher's exact test).

Study	Stimuli	Trials	Delay	Test	% Vis. Sel.	% Delay Sel.
Fuster and Jervey (1982)	2 or 4 colored lights	≥ 5	16-20 s	t-test	29% (123/417)	10% (41/416)
Miyashita and Chang (1988)	64 fractals	7-16	16 s	?	77% (144/188)	41% (77/188)
Eskandar, Richmond, and Optican (1992)	32 Walsh stimuli	≥ 7 ($\mu = 15$)	550 ms (last 256 ms)	ANOVA ($\alpha = 0.05$)	82% (41/55)	N.S.
Colombo and Gross (1994)	2 (face and pattern)	30-40	5 s	t-test ($\alpha = 0.01$)	77% (46/60)	18% (11/60)
Mikami (1995)	?	?	1-4 s	?	only selective	11% (19/174)
Chelazzi et al. (1998)	2 shapes eliciting differential response	120-144	1500 ms (last 500 ms)	t-test ($\alpha = 0.05$)	only selective	49% (54/111)
Freedman et al. (2003)	54 morphed cats and dogs	>10	1 s (300-1100 ms)	ANOVA ($\alpha = 0.01$)	48% (213/443)	7% (29/443)
Sobotka et al. (2005)	2 complex images	17	8-10 s (last 2 s)	t-test ($\alpha = 0.05$)	unspecified	36% (75/211)
Woloszyn and Sheinberg (2009)	2 images eliciting differential response	18	1 s (last 750 ms)	Mann-Whitney U ($\alpha = 0.05$)	only selective	42% (33/79)
Chapter 3	52 displays of colored squares	≥ 5 ($\mu = 9$)	1.2-1.5 s (300-1200 ms)	permutation ANOVA ($\alpha = 0.01$)	61% (95/155)	3% (4/155)

Table A.1: Previous Reports of IT Delay Period Selectivity. **Study** is the publication from which the data were extracted. **Stimuli** is the stimulus set used for the visual short-term memory task in the study. **Trials** is the number of repetitions per stimulus analyzed for delay period activity. **Delay** is the length of the delay period, with the analysis period relative to stimulus offset given in parentheses if it was not the entire period. **Test** is the statistical testing procedure used, including the significance level (α) if specified. **% Vis. Sel.** is the proportion of visually selective neurons among those analyzed, or "only selective" if only visually selective units were recorded in the task. **% Delay Sel.** is the proportion of neurons with significant delay period selectivity. (N.S. = not significant; no proportion was provided in this study.) In the latter two columns, the number of selective units and the total number of units are given in parentheses.

Appendix B

Prefrontal Stimulus Representation During Passive Fixation

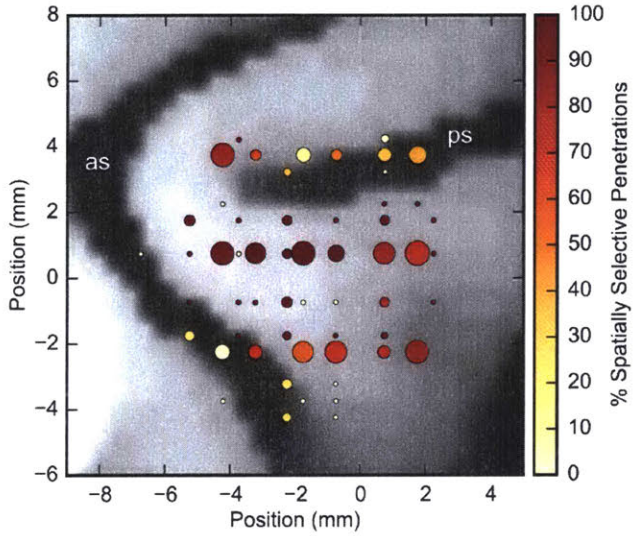
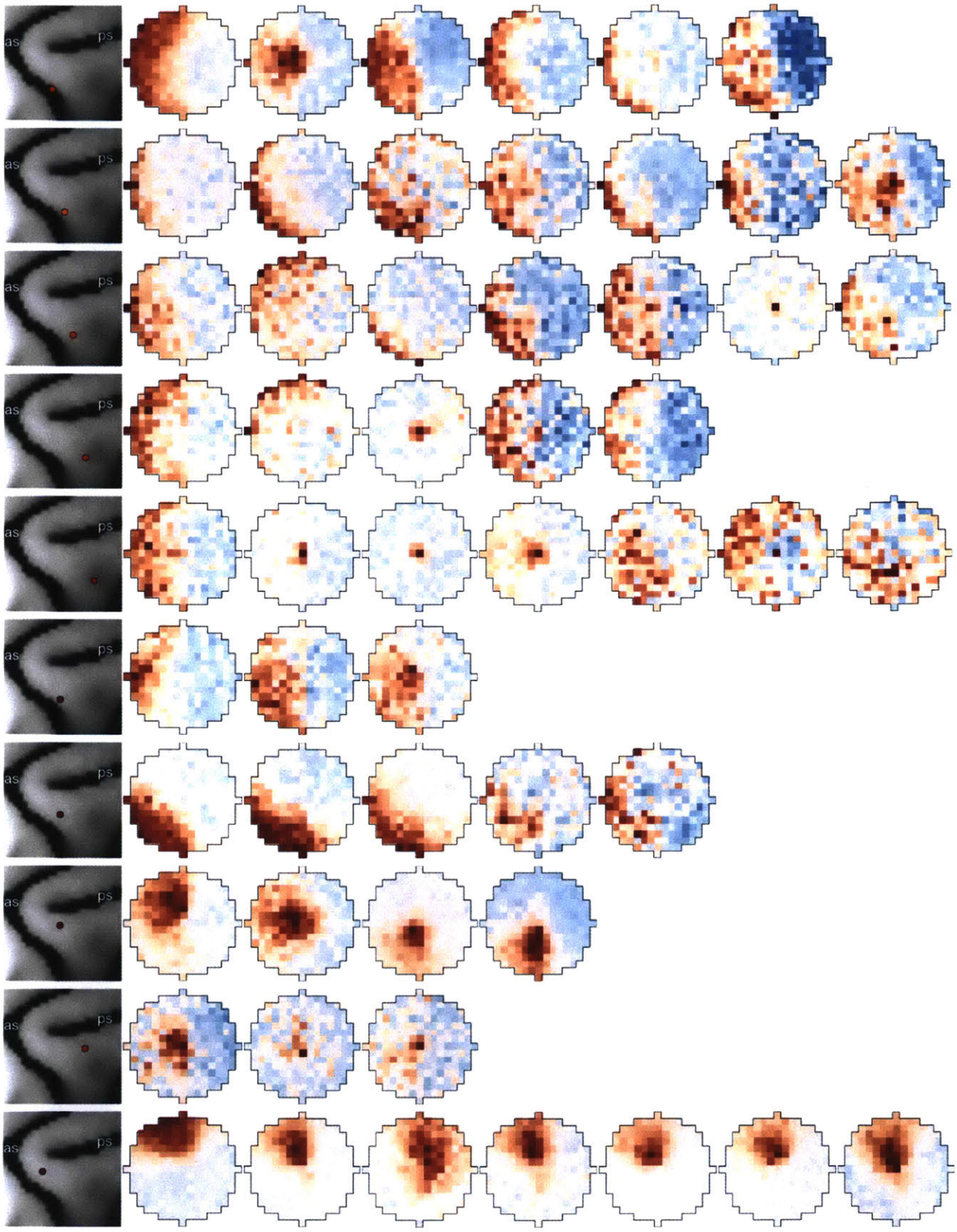
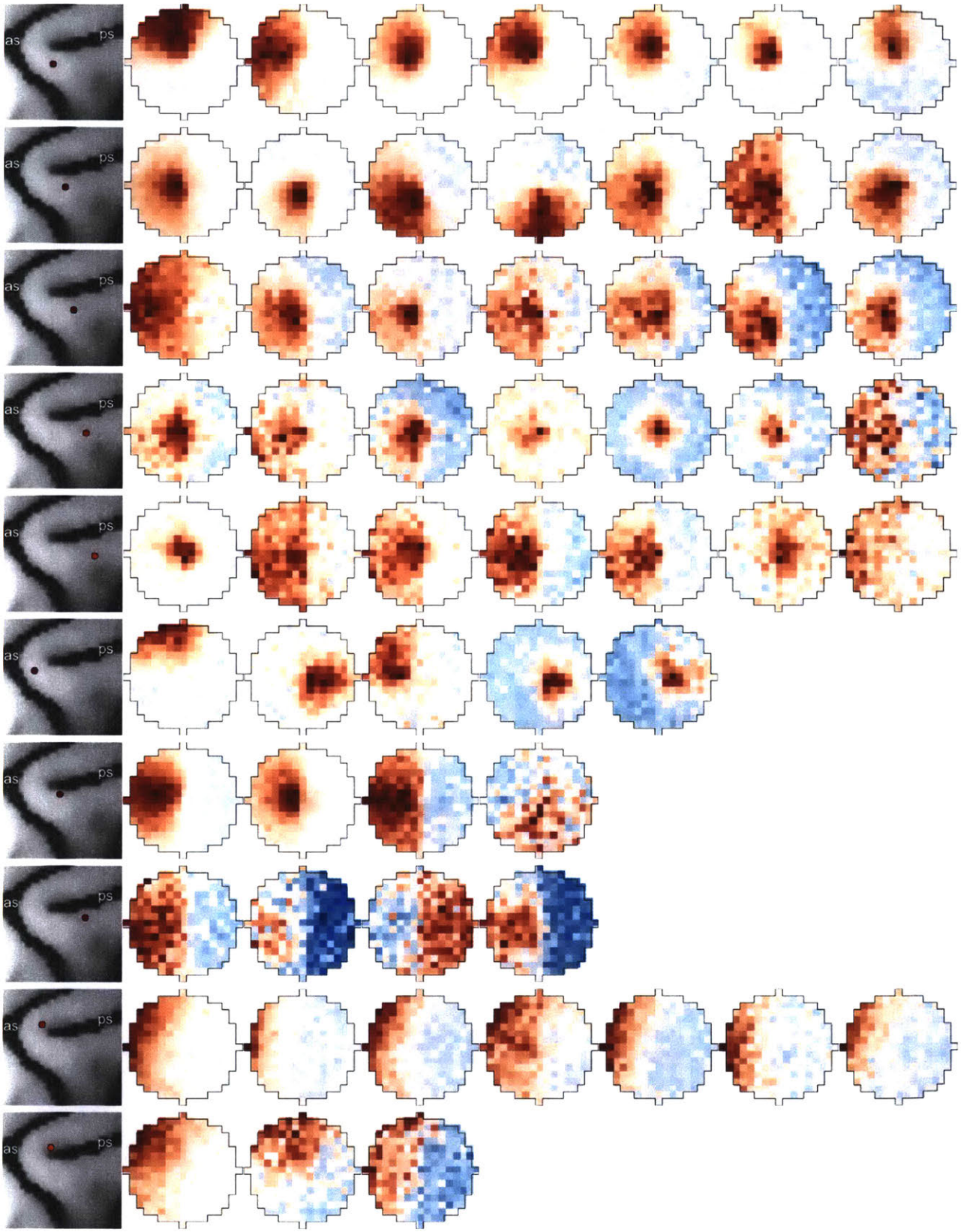


Figure B-1: Distribution of Position-Selective Activity During Passive Fixation. Proportions of penetrations for each site that yielded significant selectivity for position ($\alpha = 0.01$, permutation test of η), as assessed by mapping receptive fields. Stimuli were grids of colored dots subtending 1.5 degrees of visual angle, presented at presented at locations spanning 24 degrees of visual angle. They were presented individually as the animal passively fixated a white dot, with a presentation time of 105 ms and a 105 ms blank interval between presentations.


as
ps


as
ps

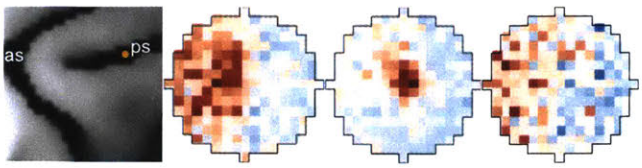


Figure B-2: Receptive Fields of Prefrontal Sites. Each row corresponds to an individual recording site in prefrontal cortex, shown at the left. Columns to the right show average multiunit activity to stimuli presented at locations spanning 24 degrees of visual angle over different recording sessions. See Figure B-1 for stimulus descriptions. Time windows were chosen to maximize the proportion of variance explained by position.

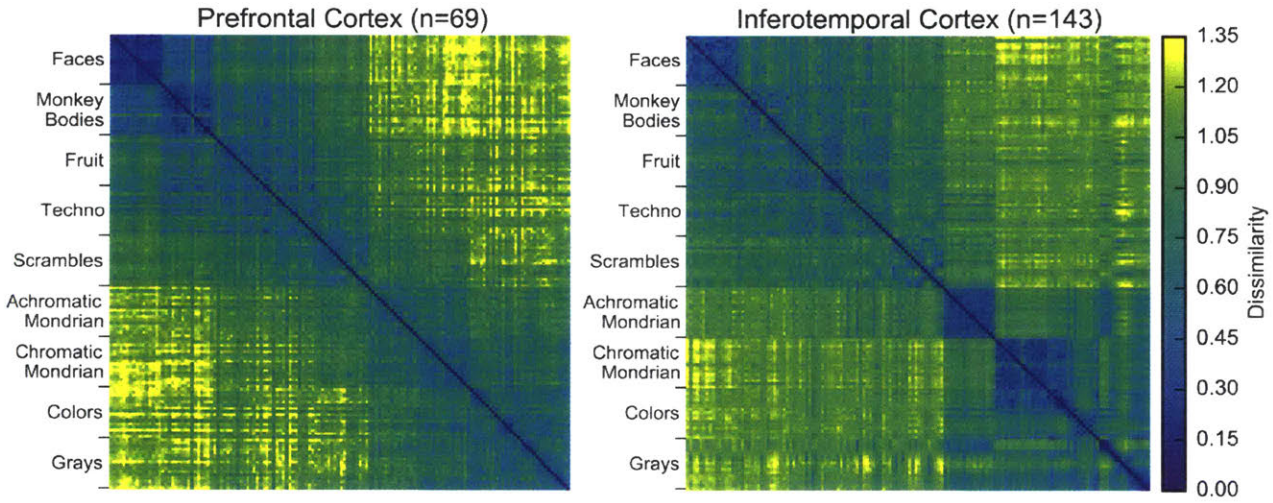


Figure B-3: Representational Dissimilarity Matrices. Representational dissimilarity matrices between population responses to different stimuli in prefrontal cortex and inferotemporal cortex. Each cell reflects an individual stimulus pair. The animal passively fixated as stimuli were presented for 105 ms, followed by a 105 ms blank period between stimuli. Activity was measured over the period from 80 to 180 ms after stimulus onset, and z-scored relative to activity during the blank period. Dissimilarity was computed as $1 - r$, where r is the Pearson correlation coefficient.

Bibliography

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111. doi: 10.1111/j.0963-7214.2004.01502006.x

Andrews, D. W. K. (1999). Estimation when a parameter is on a boundary. *Econometrica*, *67*(6), 1341–1383.

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. doi: 10.1016/j.tics.2012.05.003

Asaad, W. F., & Eskandar, E. N. (2008). A flexible software tool for temporally-precise behavioral control in Matlab. *Journal of Neuroscience Methods*, *174*(2), 245–258. doi: 10.1016/j.jneumeth.2008.07.014

Asanuma, C., Andersen, R. A., & Cowan, W. M. (1985). The thalamic relations of the caudal inferior parietal lobule and the lateral prefrontal cortex in monkeys: Divergent cortical projections from cell clusters in the medial pulvinar nucleus. *Journal of Comparative Neurology*, *241*(3), 357–381. doi: 10.1002/cne.902410309

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628. doi: 10.1111/j.1467-9280.2007.01949.x

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., & Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, 200911531. doi: 10.1073/pnas.0911531107

Balan, P. F., Oristaglio, J., Schneider, D. M., & Gottlieb, J. (2008). Neuronal correlates of the set-size effect in monkey lateral intraparietal area. *PLoS Biol*, *6*(7), e158. doi: 10.1371/journal.pbio.0060158

Barbas, H., & Zikopoulos, B. (2007). The prefrontal cortex and flexible behavior. *The Neuroscientist*, *13*(5), 532–545. doi: 10.1177/1073858407301369

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., ... Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390–401. doi: 10.1016/j.neuron.2014.12.018

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *Journal of Neuroscience*, *34*(10), 3632–3645. doi: 10.1523/JNEUROSCI.3204-13.2014

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7. doi:

10.1167/9.10.7

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, *11*(10), 6–6. doi: 10.1167/11.10.6

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851–854. doi: 10.1126/science.1158023

Bender, D. B. (1981). Retinotopic organization of macaque pulvinar. *J Neurophysiol*, *46*(3), 672–693.

Bichot, N. P., Heard, M. T., DeGennaro, E. M., & Desimone, R. (2015). A source for feature-based attention in the prefrontal cortex. *Neuron*, *88*(4), 832–844. doi: 10.1016/j.neuron.2015.10.001

Blatt, G. J., Andersen, R. A., & Stoner, G. R. (1990). Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *Journal of Comparative Neurology*, *299*(4), 421–445. doi: 10.1002/cne.902990404

Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1991). Visual topography of area TEO in the macaque. *Journal of Comparative Neurology*, *306*(4), 554–575. doi: 10.1002/cne.903060403

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences*, 201520027. doi: 10.1073/pnas.1520027113

Brody, C. D., Hernández, A., Zainos, A., & Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cerebral Cortex*, *13*(11), 1196–1207. doi: 10.1093/cercor/bhg100

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., & Miller, E. K. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, *76*(4), 838–846. doi: 10.1016/j.neuron.2012.09.029

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, *315*, 1860–1862. doi: 10.1126/science.1138071

Buschman, T. J., & Miller, E. K. (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron*, *63*(3), 386–396. doi: 10.1016/j.neuron.2009.06.020

Buschman, T. J., Siegel, M., Roy, J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences*, *108*(27), 11252–11255. doi: 10.1073/pnas.1104666108

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62. doi: 10.1038/nrn3136

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, *80*(6), 2918–2940.

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (2001). Responses of neurons in macaque area V4 during memory-guided visual search. *Cerebral Cortex*, *11*(8),

761–772. doi: 10.1093/cercor/11.8.761

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J.-D. (2017). The distributed nature of working memory. *Trends in Cognitive Sciences*, *21*(2), 111–124. doi: 10.1016/j.tics.2016.12.007

Colombo, M., & Gross, C. G. (1994). Responses of inferior temporal cortex and hippocampal neurons during delayed matching-to-sample in monkeys (Macaca fascicularis). *Behavioral Neuroscience*, *108*(3), 443–455. doi: 10.1037/0735-7044.108.3.443

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, *27*(3), 183–228. doi: 10.3758/BF03204258

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, *24*(1), 87–114.

Delacour, J. (1977). Cortex inférotemporal et mémoire visuelle à court terme chez le singe. nouvelles données. *Experimental Brain Research*, *28*(3-4), 301–310. doi: 10.1007/BF00235711

DiCarlo, J. J., & Maunsell, J. H. R. (2003). Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *Journal of Neurophysiology*, *89*(6), 3264–3278. doi: 10.1152/jn.00358.2002

Engel, A. K., & Fries, P. (2010). Beta-band oscillations — signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165. doi: 10.1016/j.conb.2010.02.015

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716. doi: 10.1038/35094565

Eskandar, E. N., Richmond, B. J., & Optican, L. M. (1992). Role of inferior temporal neurons in visual memory. I. Temporal encoding of information about visual images, recalled images, and behavioral context. *Journal of Neurophysiology*, *68*(4), 1277–1295.

Everling, S., Tinsley, C. J., Gaffan, D., & Duncan, J. (2006). Selective representation of task-relevant objects and locations in the monkey prefrontal cortex. *European Journal of Neuroscience*, *23*(8), 2197–2214. doi: 10.1111/j.1460-9568.2006.04736.x

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications*, *3*, 1229. doi: 10.1038/ncomms2237

Fox, J. (2008). *Applied regression analysis and generalized linear models.* Los Angeles: SAGE Publications.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, *23*(12), 5235–5246.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291*(5502), 312 –316. doi: 10.1126/science.291.5502.312

Fries, P., Womelsdorf, T., Oostenveld, R., & Desimone, R. (2008). The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *Journal of Neuroscience*, *28*(18), 4823–4835. doi: 10.1523/JNEUROSCI.4499-07.2008

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177–182. doi: 10.1016/j.conb.2010.03.005

Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology*, *37*, 66–74. doi: 10.1016/j.conb.2016.01.010

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, *173*(3997), 652–654. doi: 10.1126/science.173.3997.652

Fuster, J. M., Bauer, R. H., & Jervey, J. P. (1985). Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research*, *330*(2), 299–307. doi: 10.1016/0006-8993(85)90689-4

Fuster, J. M., & Jervey, J. P. (1982). Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *Journal of Neuroscience*, *2*(3), 361–375.

Gaffan, D., & Murray, E. A. (1992). Monkeys (Macaca fascicularis) with rhinal cortex ablations succeed in object discrimination learning despite 24-hr intertrial intervals and fail at matching to sample despite double sample presentations. *Behavioral Neuroscience*, *106*(1), 30–38. doi: 10.1037/0735-7044.106.1.30

Gold, J. M., Fuller, R. L., Robinson, B. M., McMahon, R. P., Braun, E. L., & Luck, S. J. (2006). Intact attentional control of working memory encoding in schizophrenia. *Journal of Abnormal Psychology*, *115*(4), 658–673. doi: 10.1037/0021-843X.115.4.658

Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *Journal of Neuroscience*, *31*(23), 8502–8511. doi: 10.1523/JNEUROSCI.0208-11.2011

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*(6213), 334–337. doi: 10.1038/338334a0

Gregoriou, G. G., Gotts, S. J., Zhou, H., & Desimone, R. (2009). High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. *Science*, *324*(5931), 1207–1210. doi: 10.1126/science.1171402

Gregoriou, G. G., Rossi, A. F., Ungerleider, L. G., & Desimone, R. (2014). Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. *Nature Neuroscience*, *17*(7), 1003–1011. doi: 10.1038/nn.3742

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194. doi: 10.1162/089892903322598139

Harada, T., Goda, N., Ogawa, T., Ito, M., Toyoda, H., Sadato, N., & Komatsu, H. (2009). Distribution of colour-selective activity in the monkey inferior temporal cortex revealed by functional magnetic resonance imaging. *European Journal of Neuroscience*, *30*(10), 1960–1970. doi: 10.1111/j.1460-9568.2009.06995.x

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632. doi: 10.1038/nature07832

Hayden, B., & Gallant, J. (2013). Working memory and decision processes in visual area

V4. *Decision Neuroscience*, *7*, 18. doi: 10.3389/fnins.2013.00018

Henderson, L. (1972). Spatial and verbal codes and the capacity of STM. *The Quarterly Journal of Experimental Psychology*, *24*(4), 485–495. doi: 10.1080/14640747208400308

Heywood, C. A., Gaffan, D., & Cowey, A. (1995). Cerebral achromatopsia in monkeys. *European Journal of Neuroscience*, *7*(5), 1064–1073. doi: 10.1111/j.1460-9568.1995.tb01093.x

Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, *69*(2), 387–396. doi: 10.1016/j.neuron.2010.12.027

Hong, H., Yamins, D. L. K., Majaj, N. J., & DiCarlo, J. J. (2016). Explicit information for category-orthogonal object properties increases along the ventral stream. *Nature Neuroscience*, *19*(4), 613–622. doi: 10.1038/nn.4247

Horel, J. A., Pytko-Joiner, D. E., Voytko, M. L., & Salsbury, K. (1987). The performance of visual tasks while segments of the inferotemporal cortex are suppressed by cold. *Behavioural Brain Research*, *23*(1), 29–42. doi: 10.1016/0166-4328(87)90240-3

Howard, M. W., Rizzuto, D. S., Caplan, J. B., Madsen, J. R., Lisman, J., Aschenbrenner-Scheibe, R., … Kahana, M. J. (2003). Gamma oscillations correlate with working memory load in humans. *Cerebral Cortex*, *13*(12), 1369–1374. doi: 10.1093/cercor/bhg084

Huxlin, K. R., Saunders, R. C., Marchionini, D., Pham, H.-A., & Merigan, W. H. (2000). Perceptual deficits after lesions of inferotemporal cortex in macaques. *Cerebral Cortex*, *10*(7), 671–683. doi: 10.1093/cercor/10.7.671

Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, *10*(3), 626–634. doi: 10.1109/72.761722

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of Neurophysiology*, *103*(4), 1963–1968. doi: 10.1152/jn.00978.2009

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*(8), 1395–1399. doi: 10.1046/j.1460-9568.2002.01975.x

Johnson, M. K., McMahon, R. P., Robinson, B. M., Harvey, A. N., Hahn, B., Leonard, C. J., … Gold, J. M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology*, *27*(2), 220–229. doi: 10.1037/a0032060

Kadohisa, M., Kusunoki, M., Petrov, P., Sigala, N., Buckley, M. J., Gaffan, D., & Duncan, J. (2015). Spatial and temporal distribution of visual information coding in lateral prefrontal cortex. *European Journal of Neuroscience*, *41*(1), 89–96. doi: 10.1111/ejn.12754

Komatsu, H., Ideura, Y., Kaji, S., & Yamane, S. (1992). Color selectivity of neurons in the inferior temporal cortex of the awake macaque monkey. *Journal of Neuroscience*, *12*(2), 408–424.

Kornblith, S., Buschman, T. J., & Miller, E. K. (2016). Stimulus load and oscillatory activity in higher cortex. *Cerebral Cortex*, *26*(9), 3772–3784. doi: 10.1093/cercor/bhv182

Kovner, R., & Stamm, J. S. (1972). Disruption of short-term visual memory by electrical stimulation of inferotemporal cortex in the monkey. *Journal of Comparative and Physiological Psychology*, *81*(1), 163–172. doi: 10.1037/h0033321

Kowalska, D. M., Bachevalier, J., & Mishkin, M. (1991). The role of the inferior prefrontal convexity in performance of delayed nonmatching-to-sample. *Neuropsychologia*, *29*(6), 583–600. doi: 10.1016/0028-3932(91)90012-W

Lafer-Sousa, R., & Conway, B. R. (2013). Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nature Neuroscience*, *16*(12), 1870–1878. doi: 10.1038/nn.3555

Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, *43*(2), 149–164. doi: 10.1016/S0042-6989(02)00402-9

Lara, A. H., & Wallis, J. D. (2014). Executive control processes underlying multi-item working memory. *Nature Neuroscience*, *17*(6), 876–883. doi: 10.1038/nn.3702

Lee, H., Simpson, G. V., Logothetis, N. K., & Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron*, *45*(1), 147–156. doi: 10.1016/j.neuron.2004.12.025

Li, C.-S. R., Mazzoni, P., & Andersen, R. A. (1999). Effect of reversible inactivation of macaque lateral intraparietal area on visual and memory saccades. *Journal of Neurophysiology*, *81*(4), 1827–1838.

Linden, D. E., Bittner, R. A., Muckli, L., Waltz, J. A., Kriegeskorte, N., Goebel, R., ... Munk, M. H. (2003). Cortical capacity constraints for visual working memory: Dissociation of fMRI load effects in a fronto-parietal network. *NeuroImage*, *20*(3), 1518–1530. doi: 10.1016/j.neuroimage.2003.07.021

Lisman, J. E., & Idiart, M. A. P. (1995). Storage of 7 +- 2 short-term memories in oscillatory subcycles. *Science*, *267*(5203), 1512–1515.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150. doi: 10.1038/35084005

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–280.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. doi: 10.1016/j.tics.2013.06.006

Lundqvist, M., Herman, P., & Lansner, A. (2011). Theta and gamma power increases and alpha/beta power decreases with memory load in an attractor network model. *Journal of Cognitive Neuroscience*, *23*(10), 3008–3020. doi: 10.1162/jocn_a_00029

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron*, *90*(1), 152–164. doi: 10.1016/j.neuron.2016.02.028

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639. doi: 10.1016/j.neuropsychologia.2010.11.031

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. doi: 10.1038/nn.3655

Maier, A., Wilke, M., Aura, C., Zhu, C., Ye, F. Q., & Leopold, D. A. (2008). Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nature neuroscience*, *11*(10), 1193–1200. doi: 10.1038/nn.2173

Makovski, T. (2012). Are multiple visual short-term memory storages necessary to explain the retro-cue effect? *Psychonomic Bulletin & Review*, *19*(3), 470–476. doi: 10.3758/s13423-012-0235-9

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 369–380. doi: 10.1037/0278-7393.34.2.369

Matsushima, A., & Tanaka, M. (2014). Different neuronal computations of spatial working memory for multiple locations within versus across visual hemifields. *Journal of Neuroscience*, *34*(16), 5621–5626. doi: 10.1523/JNEUROSCI.0295-14.2014

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex*, *43*(1), 77–94. doi: 10.1016/S0010-9452(08)70447-7

McKee, J. L., Riesenhuber, M., Miller, E. K., & Freedman, D. J. (2014). Task dependence of visual and category representations in prefrontal and inferior temporal cortices. *Journal of Neuroscience*, *34*(48), 16065–16075. doi: 10.1523/JNEUROSCI.1660-14.2014

Meltzer, J. A., Zaveri, H. P., Goncharova, I. I., Distasio, M. M., Papademetris, X., Spencer, S. S., ... Constable, R. T. (2008). Effects of working memory load on oscillatory power in human intracranial EEG. *Cerebral Cortex*, *18*(8), 1843–1855. doi: 10.1093/cercor/bhm213

Meyer, T., Qi, X.-L., Stanford, T. R., & Constantinidis, C. (2011). Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *Journal of Neuroscience*, *31*(17), 6266–6276. doi: 10.1523/JNEUROSCI.6798-10.2011

Meyers, E. M., Qi, X.-L., & Constantinidis, C. (2012). Incorporation of new information into prefrontal cortical activity after learning working memory tasks. *Proceedings of the National Academy of Sciences*, *109*(12), 4651–4656. doi: 10.1073/pnas.1201022109

Mikami, A. (1995). Visual neurons with higher selectivity can retain memory in the monkey temporal cortex. *Neuroscience Letters*, *192*(3), 157–160. doi: 10.1016/0304-3940(95)11634-9

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. doi: 10.1146/annurev.neuro.24.1.167

Miller, E. K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science*, *263*(5146), 520–522. doi: 10.1126/science.8290960

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual work-

ing memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*(16), 5154–5167.

Miller, E. K., Gochin, P. M., & Gross, C. G. (1991). Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque. *Visual Neuroscience*, *7*(04), 357–362. doi: 10.1017/S0952523800004843

Miller, E. K., Gochin, P. M., & Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Research*, *616*(1–2), 25–29. doi: 10.1016/0006-8993(93)90187-R

Miller, E. K., Li, L., & Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience*, *13*(4), 1460–1478.

Miller, G. A. (1956). The magic number seven, plus or minus two. *Psychological Review*, *63*(2).

Mirpour, K., & Bisley, J. W. (2012). Dissociating activity in the lateral intraparietal area from value using a visual foraging task. *Proceedings of the National Academy of Sciences*, *109*(25), 10083–10088. doi: 10.1073/pnas.1120763109

Missal, M., Vogels, R., Li, C.-Y., & Orban, G. A. (1999). Shape interactions in macaque inferior temporal neurons. *Journal of Neurophysiology*, *82*(1), 131–142.

Mitchell, D. J., & Cusack, R. (2011). The temporal evolution of electromagnetic markers sensitive to the capacity limits of visual short-term memory. *Frontiers in Human Neuroscience*, *5*, 18. doi: 10.3389/fnhum.2011.00018

Mittlböck, M., & Heinzl, H. (2002). Measures of explained variation in gamma regression models. *Communications in Statistics - Simulation and Computation*, *31*(1), 61–73. doi: 10.1081/SAC-9687282

Miyashita, Y., & Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, *331*(6151), 68–70. doi: 10.1038/331068a0

Moeller, S., Crapse, T., Chang, L., & Tsao, D. Y. (2017). The effect of face patch microstimulation on perception of faces and objects. *Nature Neuroscience*, *advance online publication*. doi: 10.1038/nn.4527

Mohler, C. W., Goldberg, M. E., & Wurtz, R. H. (1973). Visual receptive fields of frontal eye field neurons. *Brain Research*, *61*, 385–389. doi: 10.1016/0006-8993(73)90543-X

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869), 1543–1546. doi: 10.1126/science.1150769

Mulliken, G. H., Bichot, N. P., Ghadooshahy, A., Sharma, J., Kornblith, S., Philcock, M., & Desimone, R. (2015). Custom-fit radiolucent cranial implants for neurophysiological recording and stimulation. *Journal of Neuroscience Methods*, *241*, 146–154. doi: 10.1016/j.jneumeth.2014.12.011

Niki, H., & Watanabe, M. (1976). Prefrontal unit activity and delayed response: Relation to cue location versus direction of response. *Brain research*, *105*(1), 79–88.

Okada, K. (2013). Is omega squared less biased? A comparison of three major effect size indices in one-way anova. *Behaviormetrika*, *40*(2), 129–147. doi: 10.2333/bhmk.40.129

Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2010). Divisive normalization in olfactory

population codes. *Neuron*, *66*(2), 287–299. doi: 10.1016/j.neuron.2010.04.009

Op De Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *Journal of Comparative Neurology*, *426*(4), 505–518. doi: 10.1002/1096-9861(20001030)426:4<505::AID-CNE1>3.0.CO;2-M

Pachitariu, M., Steinmetz, N. A., Kadir, S. N., Carandini, M., & Harris, K. D. (2016). Fast and accurate spike sorting of high-channel count probes with KiloSort. *Advances in Neural Information Processing Systems*, *29*, 4448–4456.

Pagan, M., Urban, L. S., Wohl, M. P., & Rust, N. C. (2013). Signals in inferotemporal and perirhinal cortex suggest an untangling of visual target information. *Nature Neuroscience*, *16*(8), 1132–1139. doi: 10.1038/nn.3433

Palva, J. M., Monto, S., Kulashekhar, S., & Palva, S. (2010). Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proceedings of the National Academy of Sciences*, *107*(16), 7580–7585. doi: 10.1073/pnas.0913113107

Palva, S., Kulashekhar, S., Hämäläinen, M., & Palva, J. M. (2011). Localization of cortical phase and amplitude dynamics during visual working memory encoding and retention. *Journal of Neuroscience*, *31*(13), 5013–5025. doi: 10.1523/JNEUROSCI.5592-10.2011

Parvizi, J., Jacques, C., Foster, B. L., Withoft, N., Rangarajan, V., Weiner, K. S., & Grill-Spector, K. (2012). Electrical stimulation of human fusiform face-selective regions distorts face perception. *Journal of Neuroscience*, *32*(43), 14915–14920. doi: 10.1523/JNEUROSCI.2609-12.2012

Passingham, R. (1975). Delayed matching after selective prefrontal lesions in monkeys (Macaca mulatta). *Brain Research*, *92*(1), 89–102. doi: 10.1016/0006-8993(75)90529-6

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*(2), 97–107. doi: 10.1038/nrn1603

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pereira, J., & Wang, X.-J. (2014). A tradeoff between accuracy and flexibility in a working memory circuit endowed with slow feedback mechanisms. *Cerebral Cortex*, bhu202. doi: 10.1093/cercor/bhu202

Petrides, M. (2000). Dissociable roles of mid-dorsolateral prefrontal and anterior inferotemporal cortex in visual working memory. *Journal of Neuroscience*, *20*(19), 7496–7503.

Politis, D. N., Romano, J. P., & Wolf, M. (1999). *Subsampling*. New York, NY: Springer New York. (OCLC: 853257811)

Postle, B. R. (2016). How does the brain keep information "in mind"? *Current Directions in Psychological Science*, *25*(3), 151–156. doi: 10.1177/0963721416643063

Rainer, G., Asaad, W. F., & Miller, E. K. (1998). Memory fields of neurons in the primate prefrontal cortex. *Proceedings of the National Academy of Sciences*, *95*(25), 15008–15013.

Rainer, G., & Miller, E. K. (2002). Timecourse of object-related neural activity in the primate prefrontal cortex during a short-term memory task. *European Journal of*

Neuroscience, *15*(7), 1244–1254. doi: 10.1046/j.1460-9568.2002.01958.x

Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, *276*(5313), 821–824. doi: 10.1126/science.276.5313.821

Ray, S., & Maunsell, J. H. (2010). Differences in gamma frequencies across visual cortex restrict their possible use in computation. *Neuron*, *67*(5), 885–896. doi: 16/j.neuron .2010.08.004

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*(2), 168–185. doi: 10.1016/j.neuron.2009.01.002

Richter, C. G., Thompson, W. H., Bosman, C. A., & Fries, P. (2015). A jackknife approach to quantifying single-trial correlation between covariance-based metrics undefined on a single-trial basis. *NeuroImage*, *114*, 57–70. doi: 10.1016/j.neuroimage.2015.04.040

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 585–590. doi: 10.1038/nature12160

Riley, M. R., Qi, X.-L., & Constantinidis, C. (2016). Functional specialization of areas along the anterior–posterior axis of the primate prefrontal cortex. *Cerebral Cortex*. doi: 10.1093/cercor/bhw190

Robitaille, N., Marois, R., Todd, J., Grimault, S., Cheyne, D., & Jolicœur, P. (2010). Distinguishing between lateralized and nonlateralized brain activity associated with visual short-term memory: fMRI, MEG, and EEG evidence from the same observers. *NeuroImage*, *53*(4), 1334–1345. doi: 10.1016/j.neuroimage.2010.07.027

Rolls, E. T., & Tovee, M. J. (1995). The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. *Experimental Brain Research*, *103*(3), 409–420. doi: 10.1007/BF00241500

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136–1139. doi: 10.1126/science.aah7011

Roux, F., Wibral, M., Mohr, H. M., Singer, W., & Uhlhaas, P. J. (2012). Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory. *Journal of Neuroscience*, *32*(36), 12411–12420. doi: 10.1523/ JNEUROSCI.0421-12.2012

Rushworth, M. F. S., Nixon, P. D., Eacott, M. J., & Passingham, R. E. (1997). Ventral prefrontal cortex is not essential for working memory. *The Journal of Neuroscience*, *17*(12), 4829–4838.

Saalmann, Y. B., Pinsk, M. A., Wang, L., Li, X., & Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. *Science*, *337*(6095), 753–756. doi: 10.1126/science.1223082

Salazar, R. F., Dotson, N. M., Bressler, S. L., & Gray, C. M. (2012). Content-specific fronto-parietal synchronization during visual working memory. *Science*, *338*(6110), 1097–1100. doi: 10.1126/science.1224000

Saleem, K. S., Miller, B., & Price, J. L. (2014). Subdivisions and connectional networks of the lateral prefrontal cortex in the macaque monkey. *Journal of Comparative Neurology*,

522(7), 1641–1690. doi: 10.1002/cne.23498

Sato, T. (1989). Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Experimental Brain Research*, *77*(1), 23–30. doi: 10.1007/BF00250563

Sawaguchi, T., & Goldman-Rakic, P. S. (1991). D1 dopamine receptors in prefrontal cortex: Involvement in working memory. *Science*, *251*(4996), 947–950. doi: 10.1126/science.1825731

Sawamura, H., Orban, G. A., & Vogels, R. (2006). Selectivity of neuronal adaptation does not match response selectivity: A single-cell study of the fMRI adaptation paradigm. *Neuron*, *49*(2), 307–318. doi: 10.1016/j.neuron.2005.11.028

Scalaidhe, S. P. Ó., Wilson, F. A. W., & Goldman-Rakic, P. S. (1999). Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: Evidence for intrinsic specialization of neuronal coding. *Cerebral Cortex*, *9*(5), 459–475. doi: 10.1093/cercor/9.5.459

Schall, J. D., Morel, A., King, D. J., & Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: Convergence and segregation of processing streams. *Journal of Neuroscience*, *15*(6), 4464–4487.

Schneegans, S., & Bays, P. M. (2016). No fixed item limit in visuospatial working memory. *Cortex*, *83*, 181–193. doi: 10.1016/j.cortex.2016.07.021

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*(2), 207–214. doi: 10.1111/j.1467-9280.2009.02276.x

Shao, J., & Wu, C. F. J. (1989). A general theory for jackknife variance estimation. *The Annals of Statistics*, *17*(3), 1176–1197.

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310. doi: 10.1214/10-STS330

Siegel, M., Donner, T. H., Oostenveld, R., Fries, P., & Engel, A. K. (2008). Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron*, *60*(4), 709–719. doi: 10.1016/j.neuron.2008.09.010

Siegel, M., Warden, M. R., & Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*, *106*(50), 21341–21346.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20. doi: 10.1016/j.tics.2004.11.006

Sirotin, Y. B., & Das, A. (2009). Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature*, *457*(7228), 475–479. doi: 10.1038/nature07664

Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLOS ONE*, *3*(2), e1699. doi: 10.1371/journal.pone.0001699

Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2009). V4 activity predicts the strength of visual short-term memory representations. *The Journal of Neuroscience*, *29*(23), 7432–7438.

Sligte, I. G., Wokke, M. E., Tesselaar, J. P., Steven Scholte, H., & Lamme, V. A. F. (2011). Magnetic stimulation of the dorsolateral prefrontal cortex dissociates fragile visual short-term memory from visual working memory. *Neuropsychologia*, *49*(6), 1578–1588. doi: 10.1016/j.neuropsychologia.2010.12.010

Sobotka, S., Diltz, M. D., & Ringo, J. L. (2005). Can delay-period activity explain working memory? *Journal of Neurophysiology*, *93*(1), 128–136. doi: 10.1152/jn.01002.2003

Sprague, T. C., Ester, E. F., & Serences, J. T. (2014). Reconstructions of information in visual spatial working memory degrade with memory load. *Current Biology*, *24*(18), 2174–2180. doi: 10.1016/j.cub.2014.07.066

Stanton, G. B., Bruce, C. J., & Goldberg, M. E. (1995). Topography of projections to posterior cortical areas from the macaque frontal eye fields. *Journal of Comparative Neurology*, *353*(2), 291–305. doi: 10.1002/cne.903530210

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405. doi: 10.1016/j.tics .2015.05.004

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, *78*(2), 364–375. doi: 10.1016/ j.neuron.2013.01.039

Supèr, H., Spekreijse, H., & Lamme, V. A. F. (2001). A neural correlate of working memory in the monkey primary visual cortex. *Science*, *293*(5527), 120–124. doi: 10.1126/ science.1060496

Suzuki, H., & Azuma, M. (1983). Topographic studies on visual neurons in the dorsolateral prefrontal cortex of the monkey. *Experimental Brain Research*, *53*(1), 47–58. doi: 10.1007/BF00239397

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*(6984), 751. doi: 10.1038/nature02466

Todd, J. J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(2), 144–155. doi: 10.3758/CABN.5.2.144

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78.

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural limits to representing objects still within view. *The Journal of Neuroscience*, *33*(19), 8257–8263. doi: 10 .1523/JNEUROSCI.5348-12.2013

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, *121*(1), 124–149. doi: 10.1037/a0035234

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*(22), 8780–8785. doi: 10.1073/pnas.1117465109

Vandenbroucke, A. R., Sligte, I. G., & Lamme, V. A. (2011). Manipulations of attention dissociate fragile visual short-term memory from visual working memory. *Neuropsychologia*, *49*(6), 1559–1568. doi: 10.1016/j.neuropsychologia.2010.12.044

van Kerkoerle, T., Self, M. W., & Roelfsema, P. R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nature Communications*, *8*, 13804. doi: 10.1038/ncomms13804

van Vugt, M. K., Schulze-Bonhage, A., Litt, B., Brandt, A., & Kahana, M. J. (2010). Hippocampal gamma oscillations increase with memory load. *Journal of Neuroscience*, *30*(7), 2694–2699. doi: 10.1523/JNEUROSCI.0567-09.2010

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., & Pennartz, C. M. (2010). The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization. *NeuroImage*, *51*(1), 112–122. doi: 10.1016/j.neuroimage.2010.01.073

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. doi: 10.1038/nature04171

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92–114. doi: 10.1037/0096-1523.27.1.92

Vogels, R. (2016). Sources of adaptation of inferior temporal cortical responses. *Cortex*, *80*, 185–195. doi: 10.1016/j.cortex.2015.08.024

Vogels, R., & Orban, G. A. (1994). Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. *Journal of Neurophysiology*, *71*(4), 1428–1451.

Vogels, R., Sáry, G., & Orban, G. A. (1995). How task-related are the responses of inferior temporal neurons? *Visual Neuroscience*, *12*(2), 207–214. doi: 10.1017/S0952523800007884

Voytek, B., & Knight, R. T. (2010). Prefrontal cortex and basal ganglia contributions to visual working memory. *Proceedings of the National Academy of Sciences*, *107*(42), 18167–18172. doi: 10.1073/pnas.1007277107

Warden, M. R., & Miller, E. K. (2007). The representation of multiple objects in prefrontal neuronal delay activity. *Cerebral Cortex*, *17*(suppl 1), i41–i50. doi: 10.1093/cercor/bhm070

Warden, M. R., & Miller, E. K. (2010). Task-dependent changes in short-term memory in the prefrontal cortex. *Journal of Neuroscience*, *30*(47), 15801–15810. doi: 10.1523/JNEUROSCI.1569-10.2010

Watanabe, K., & Funahashi, S. (2014). Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nature Neuroscience*, *17*(4), 601–611. doi: 10.1038/nn.3667

Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cerebral Cortex*, *4*(5), 470–483. doi: 10.1093/cercor/4.5.470

Wei, Z., Wang, X.-J., & Wang, D.-H. (2012). From distributed resources to limited slots in multiple-item working memory: A spiking network model with normalization. *Journal*

of Neuroscience, *32*(33), 11228–11240. doi: 10.1523/JNEUROSCI.0735-12.2012

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64. doi: 10.1037/0096-3445.131.1.48

Wilson, F. A. W., Scalaidhe, S. P. Ó., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*, *260*(5116), 1955–1958.

Wilson, N. R., Runyan, C. A., Wang, F. L., & Sur, M. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature*, *488*(7411), 343–348. doi: 10.1038/nature11347

Woloszyn, L., & Sheinberg, D. L. (2009). Neural dynamics in inferior temporal cortex during a visual working memory task. *Journal of Neuroscience*, *29*(17), 5494–5507. doi: 10.1523/JNEUROSCI.5785-08.2009

Womelsdorf, T., Fries, P., Mitra, P. P., & Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature*, *439*(7077), 733–736. doi: 10.1038/nature04258

Xu, Y. (2007). The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. *The Journal of Neuroscience*, *27*(43), 11676–11686. doi: 10.1523/JNEUROSCI.3545-07.2007

Xu, Y. (2008). Distinctive neural mechanisms supporting visual object individuation and identification. *Journal of Cognitive Neuroscience*, *21*(3), 511–518. doi: 10.1162/jocn.2008.21024

Xu, Y., & Chun, M. M. (2005). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95. doi: 10.1038/nature04262

Yang, Y., Ming, J., & Yu, N. (2012). Color image quality assessment based on CIEDE2000. *Advances in Multimedia*, *2012*, e273723. doi: 10.1155/2012/273723

Yasuda, M., Banno, T., & Komatsu, H. (2010). Color selectivity of neurons in the posterior inferior temporal cortex of the macaque monkey. *Cerebral Cortex*, *20*(7), 1630–1646. doi: 10.1093/cercor/bhp227

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233. doi: 10.1038/nature06860

Zoccolan, D., Cox, D. D., & DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *Journal of Neuroscience*, *25*(36), 8150–8164. doi: 10.1523/JNEUROSCI.2058-05.2005